NPS ARCHIVE
1997.12
GORDON, K.

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A METRIC EVALUATION APPROACH FOR THE DEFENSE ACQUISITION WORKFORCE IMPROVEMENT ACT

by

Kathy E. Gordon

December 1997

Thesis Advisor: Walter E. Owen

Approved for public release; distribution is unlimited.

Thesis
G574325

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE December 1997	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE A METRIC EVALUATION APPROACH FOR THE DEFENSE ACQUISITION WORKFORCE MPROVEMENT ACT

5. FUNDING NUMBERS

6. AUTHOR(S) Gordon, Kathy E.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

The Defense Acquisition Workforce Improvement Act (DAWIA) was enacted in 1990 to improve the quality and professionalism of the Department of Defense acquisition workforce. To assess the effectiveness of DAWIA, actual outcomes of the law must be compared to its original objectives. A particular difficulty in evaluating public policies is that they usually cannot be measured in terms of output, such as dollars, hours, or physical units. The primary objective of this study was to find and introduce a performance measurement approach suitable for identifying effective metrics. A second objective was to establish a link between metrics and outcomes. Using the performance measurement approach as a tool, an analysis attempted to link acquisition workforce metrics with specific outcomes. To explore this issue, a literature review of relevant organizational and management texts on public policy analysis, performance measurement and program evaluation was conducted. Four suitable frameworks for performance measurement were found and evaluated. The preferred approach for evaluating DAWIA was determined to be a combination of two performance measurement approaches. The new approach is called Metric Assessment and Measurement Approach. It includes valuable features of the two approaches, and a newly developed metric template for evaluating metrics.

14. SUBJECT TERMS
Performance Measurement, Metrics, DAWIA

15. NUMBER OF PAGES 94

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited.

A METRIC EVALUATION APPROACH FOR THE DEFENSE ACQUISITION WORKFORCE IMPROVEMENT ACT

Kathy E. Gordon
Lieutenant Commander, United States Navy
B.S., Ball State University, 1981

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 1997

NPS ARCHIVE
1997.12
GORDON, K.

X1051
G74325
C.2

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

ABSTRACT

The Defense Acquisition Workforce Improvement Act (DAWIA) was enacted in 1990 to improve the quality and professionalism of the Department of Defense acquisition workforce. To assess the effectiveness of DAWIA, actual outcomes of the law must be compared to its original objectives. A particular difficulty in evaluating public policies is that they usually cannot be measured in terms of output, such as dollars, hours, or physical units. The primary objective of this study was to find and introduce a performance measurement approach suitable for identifying effective metrics. A second objective was to establish a link between metrics and outcomes. Using the performance measurement approach as a tool, an analysis attempted to link acquisition workforce metrics with specific outcomes. To explore this issue, a literature review of relevant organizational and management texts on public policy analysis, performance measurement and program evaluation was conducted. Four suitable frameworks for performance measurement were found and evaluated. The preferred approach for evaluating DAWIA was determined to be a combination of two performance measurement approaches. The new approach is called Metric Assessment and Measurement Approach. It includes valuable features of the two approaches, and a newly developed metric template for evaluating metrics.

TABLE OF CONTENTS

I. INTRODUCTION 1

A. AREA OF RESEARCH 1

B. BACKGROUND 2

C. RESEARCH QUESTIONS 3

D. SCOPE OF THESIS 4

E. METHODOLOGY 4

F. ORGANIZATION 5

G. BENEFITS OF STUDY 6

II. LITERATURE REVIEW 7

A. INTRODUCTION 7

B. HISTORY OF ACQUISITION REFORM 7

1. Hoover Commission 9

2. Grace Commission 10

3. Packard Commission 10

4. Institute for Defense Analyses Study 11

C. DEFENSE ACQUISITION WORKFORCE IMPROVEMENT ACT . . 12

D. WHY MEASURE PERFORMANCE? 14

E. PERFORMANCE MEASUREMENT THEORY 16

1. Categories of Measurement 17

a. Validity 19
b. Reliability 19
c. Bias and Sampling Errors 20
d. Judgmental vs. Nonjudgmental Measures 21
3. Measurement Paradigms 22
F. METRICS 24
1. Definition of Metrics 24
2. Attributes of Metrics 25
3. Metric Development 26
4. Classification of Metrics 27
5. Examples of Metrics 28
G. CHAPTER SUMMARY 29

III. PERFORMANCE MEASUREMENT APPROACHES 31
A. INTRODUCTION 31
B. ALTERNATIVE APPROACHES 32
1. General Measurement Methodology 33
2. Performance Measurement Process 36
3. Measurement Linkage Model™ 38
4. Metric Development Process 43
C. APPROACH SELECTED 51
D. CHAPTER SUMMARY 53

IV. DATA ANALYSIS 55

A. INTRODUCTION 55

B. METRIC ASSESSMENT AND MEASUREMENT APPROACH 55

1. Metric Template 59

2. Metric Package 61

C. VALIDATION OF APPROACH 64

D. CHAPTER SUMMARY 66

V. CONCLUSIONS AND RECOMMENDATIONS 69

A. CONCLUSIONS 69

B. RESEARCH QUESTIONS ANSWERED 70

C. RECOMMENDATIONS DRAWN FROM RESEARCH 71

D. RECOMMENDATIONS FOR FURTHER RESEARCH 72

APPENDIX 75

LIST OF REFERENCES 77

BIBLIOGRAPHY 81

INITIAL DISTRIBUTION LIST 83

I. INTRODUCTION

A. AREA OF RESEARCH

In assessing the effectiveness of public law or policy, the primary question to be answered is: did it achieve the desired outcome? To answer this question, actual consequences of a policy or law must be compared to their objectives. It is often believed that when Congress passes new legislation, the ensuing results will be what the law was intended to achieve. "Unfortunately, these assumptions are not always warranted. The national experiences with many public programs indicate the need for careful appraisal of the real impact of public policy" (Dye, 1995, p. 320). Cave, Kogan and Smith (1990) argued that "there are few people in the public sector who would disagree with the proposition that, if objectives are clear, it is helpful to compare actual performance with intended performance. Such comparisons enable effort to be redirected if performance is not what it should be and help to increase the probability of ultimate success" (p. 10).

A particular difficulty in evaluating public policy is that it usually cannot be measured in terms of output, such as dollars or units. Public policies result in outcomes, not outputs. Approaches most often used to evaluate public policy are:

- congressional hearings or reports
- site visits
- program measures or statistics
- application of professional standards
- review of citizen complaints (Dye, 1995).

This study will attempt to identify useful program measures to evaluate public policy. To be useful, measures must impart information about outcomes. Consequently, this research attempts to link measures (also called "metrics") to outcomes. These measures must be accurate indicators of performance. Accurate measures will enable analysts to evaluate the effectiveness of public policy, and provide an answer to the question "did it achieve the desired outcome?"

In the Department of Defense (DoD), many are asking if the Defense Acquisition Workforce Improvement Act (DAWIA) has achieved its desired outcomes. This Act is one of many defense acquisition reforms intended to result in a more effective and better trained acquisition workforce, which in turn, is expected to save taxpayer money by improving acquisition of military hardware.

In addition to accurate and effective measures, a formal process must be used to evaluate DAWIA to determine if it has achieved its desired outcomes. To do this, four performance measurement approaches will be introduced that are suitable for evaluating performance. Using the most suitable approach as a tool, an analysis will attempt to link acquisition workforce metrics with specific outcomes.

B. BACKGROUND

About 110,000 federal workers, both civilian and military, are employed in one of 11 different career fields that make up the defense acquisition workforce. They are responsible for procuring goods and services for the Defense Department. The defense acquisition structure is described as "a single uniform system whereby all equipment, facilities, and services are planned, developed, acquired, maintained, and disposed of by the Department of

Defense" (Schmoll, 1996, p. 1).

The Defense Acquisition Workforce Improvement Act is a federal statute enacted in 1990 (and amended several times since) in response to the need for better management of the defense acquisition process. The purpose of DAWIA is to improve acquisition practices in the Defense Department through management of the acquisition workforce. According to the U.S. General Accounting Office, "Specific provisions of the act require the Secretary of Defense to (1) establish a management structure and policies and regulations for implementing the act's provisions, (2) establish qualification requirements, (3) provide training and education to meet these requirements, and (4) enhance civilian opportunities to progress to senior acquisition positions" (GAO (i), 1993, p. 1).

The act aims to provide greater stability to acquisition programs through reduction in workforce turnover by means of managed education, training, and career development of the acquisition workforce. The ultimate objective is better control of the acquisition process through a higher quality and better trained workforce.

C. RESEARCH QUESTIONS

The primary research question is: Which metrics are good indicators of acquisition workforce performance?

Subsidiary research questions are:

- How can suitable metrics for workforce performance be identified?
- How can metrics be evaluated as good indicators using a performance measurement tool?
- Can a metric template be developed that captures these performance indicators?

- How can acquisition workforce metrics be linked to performance improvement outcomes?

D. SCOPE OF THESIS

This thesis provides a brief history of acquisition reform and summarizes the Defense Acquisition Workforce Improvement Act. It discusses performance measurement theory, defines metrics, and compares and evaluates four performance measurement approaches. It proposes and defends a suitable approach for assessing and analyzing the effectiveness of DAWIA metrics. Finally, it recommends a metric template format for identifying fundamental performance measures in the workplace, and provides recommendations for follow-on research.

E. METHODOLOGY

A literature review was conducted of relevant organizational behavior and management texts, and writings on public policy analysis, performance measurement, and program evaluation. Defense publications, General Accounting Office reports, and transcripts of congressional subcommittee hearings were examined. In addition, Internet websites and homepages were queried for up-to-date defense acquisition information.

This study followed Rhoads' (1995) methodology outlined below. He suggested the following steps in analyzing the impact of DAWIA:

- define study objectives
- define problem domain and establish boundaries
- identify metrics
- develop a model

- identify data to be collected and sources of data
- collect data
- analyze and interpret data
- report the results of the study (Rhoads, 1995, p. 98)

This research will perform only the first five steps of Rhoads' methodology. The main objective of this study is to identify metrics that will enable analysts to collect, examine and interpret appropriate data on their own.

F. ORGANIZATION

This thesis consists of five chapters. The next chapter is a literature review which traces the history of acquisition reform from post World War II through enactment of DAWIA. Brief summaries of several significant defense acquisition studies are summarized in Chapter II. A description DAWIA's provisions and objectives is provided. Following that is a discussion of performance measurement theory and metric concepts.

Chapter III introduces four performance measurement approaches discovered in the course of the literature review, and discusses strengths and weaknesses of each approach. A suitable performance measurement approach is then presented and defended.

Chapter IV discusses application of the preferred approach in greater detail, and introduces a format for linking metrics to specific outcomes. The final chapter offers conclusions drawn from the analysis, and makes recommendations for further research. A glossary of frequently used terms is included in an appendix following Chapter V.

G. BENEFITS OF STUDY

The study provides a performance measurement approach that is useful to gauge the success of acquisition reform in particular areas, specifically workforce performance as a result of the Defense Acquisition Workforce Improvement Act. It may focus efforts to update or revise acquisition training material in response to workforce deficiencies. In addition, this research is expected to prioritize and improve acquisition policies and procedures by helping to identify what things or activities to measure, thereby enabling corrective actions if necessary.

II. LITERATURE REVIEW

A. INTRODUCTION

This chapter reviews the history of acquisition reform initiatives since the mid 1950s, leading up to passage of the Defense Acquisition Workforce Improvement Act in 1990. Several notable studies of defense acquisition practices were undertaken during that 35 year period, including the Hoover, Grace and Packard Commissions. The Defense Acquisition Workforce Improvement Act is a milestone in recognizing the fundamental role of a skilled workforce in improving acquisition practices. The Act's provisions and objectives are summarized in a subsequent section.

Follow-on sections discuss performance measurement theory and paradigms. The concept of metrics is explained, along with a discussion of developing and categorizing metrics. And finally, examples of metrics are given.

B. HISTORY OF ACQUISITION REFORM

The Defense Acquisition Workforce Improvement Act is one of many Department of Defense initiatives aimed at improving the way DoD acquires weapon systems. Acquisition reform has risen to the forefront of public attention in recent times. Several factors are driving this increased interest. One factor is public pressure to balance the federal budget, which has resulted in a disproportionate share of cuts in defense spending. With an ever-shrinking defense budget, ways to spend military procurement money more wisely must be explored. Recent changes to conventional acquisition practices have resulted in significant savings, both in time and cost.

Another catalyst for acquisition reform is the public's impression of reckless government spending. Congress found that "waste and inefficiency in Federal programs undermine the confidence of the American people in the Government and reduces the Federal Government's ability to address adequately vital public needs" (GPRA, 1993, p. 1). In the public's eyes, the Department of Defense is among the worst offenders in mismanaging taxpayer funds; whether real or perceived, this is a common impression among the public.

Within the Department of Defense, acquisition is defined as the conceptualization, development, test, production, deployment, modification, and disposal of military weapons systems. Procurement is simply the purchase of goods and services for military use. The two terms are often used interchangeably in the broader context (i.e., acquisition), as is the practice throughout this document.

There have been many initiatives over the last several decades to address military acquisition practices, but no real change in the way the government acquires its weapons systems has taken place until enactment of DAWIA. A DAWIA summary (ACT) noted that, "Although it has been recognized for over 30 years in studies and commissions (including the First and Second Hoover Commissions in 1949 and 1955, the Fitzhugh Commission in 1970, the Commission on Government Procurement in 1972, and the Packard Commission in 1986) that the quality and professionalism of the Department of Defense acquisition workforce should be improved, implementation of the recommendations of these commissions has been sorely lacking" (ACT, 1997, p. 1).

This section provides a brief history of four significant procurement studies that preceded the Defense Acquisition Workforce Improvement Act. Common to all four studies

was the recognition that real change depended on a better educated and more highly trained workforce.

1. Hoover Commission

In 1953, Congress formed the "Hoover Commission"[1] (officially known as the Commission on Organization of the Executive Branch of the Government) to look into ways to save money in defense spending. Chaired by former President Herbert Hoover, the Commission was directed to investigate government operations to find ways to:

> . . . promote economy, efficiency, and improved service in . . . the executive branch of the Government by (1) recommending methods and procedures for reducing expenditures to the lowest amount consistent with the efficient performance of essential services, activities, and functions; . . . . (COMMISSION (i), 1955, p. 23)

Regarding military procurement, the Hoover Commission reported that "great potential savings" could result from their recommendations, but gave no dollar estimate. A Commission recommendation asserted: "The Secretary of Defense should establish a policy requiring each military department to develop and assign career-trained personnel to technical and executive posts throughout the field of procurement management" (COMMISSION (ii), 1955, p. 131). Regrettably, it was four decades before this recommendation was codified into law.

---

[1]The Hoover Commission that was formed in 1953 was, in fact, the second Hoover Commission; the first Hoover Commission was concerned primarily with finding savings through government reorganization.

2. Grace Commission

President Reagan in 1982 chartered the Private Sector Survey on Cost Control, better known as the Grace Commission. W.R. Grace & Company conducted this procurement study with over 2,000 private sector volunteers at no cost to the government. Two key findings of the study were inconsistent quality of acquisition personnel, and inadequate support to Program Managers. The Grace Commission recommended consolidating all defense acquisition functions under a single senior acquisition executive, and establishing career paths and technical training for acquisition personnel.

3. Packard Commission

Former Deputy Secretary of Defense David Packard led the next large scale examination of defense acquisition practices in 1985. The Packard Commission was charged by executive order with reviewing defense acquisition within the military departments.

The Packard Commission's key findings were lack of accountability in the acquisition system, and inadequate career development and education of the defense acquisition workforce. Principal recommendations for improving the quality of acquisition personnel included:

- establishing an Under Secretary of Defense for Acquisition
- establishing professional qualifications for the acquisition workforce
- expanding education and training opportunities for acquisition personnel.

The Packard report (PACKARD) emphasized: "The Commission's recommendations in this critical area can and should be acted upon quickly and are of the highest priority" (PACKARD, 1986, p. 14). Reagan accepted the Commissions findings, and ordered its

recommendations be carried out. Many recommendations were acted upon, including establishing the post of Under Secretary of Defense for Acquisition.

4. Institute for Defense Analyses Study

A study chartered by the Under Secretary of Defense for Acquisition was conducted by the Institute for Defense Analyses (IDA) in 1988 to review how well DoD had done in implementing the Packard recommendations. The IDA report praised DoD for major organizational changes brought about by the Packard Commission, including creating the position of Under Secretary of Defense for Acquisition. However, the report also expressed veiled dissatisfaction with DoD's lack of progress in improving workforce quality. In testimony before the House Investigations Subcommittee, David Graham, IDA Project Director, highlighted DoD's shortcomings. In an exchange with Rep. Dennis Hertel (D.-Mich.), Graham's testimony underscored the need for better personnel management. Excerpts of his testimony follow:

> MR. GRAHAM: Our agenda for personnel management is that as a first step, the Secretary [of Defense] should assign a senior acquisition staff member to monitor the acquisition personnel system. We discovered that no one in the department really handles this issue; no one monitors the qualifications of the people out in the field, the contract managers, the training and career paths for program managers.
> MR. HERTEL: Mr. Chairman.
> MR. MAVROULES: Yes.
> MR. HERTEL: That is incredible. You found under this study that no one in the entire Department of Defense is responsible for all the people, the personnel who buy all the weapons with $300 billion in the total budget?
> MR. GRAHAM: For the people. No one person. That is correct. (HEARINGS (i), 1990, pp. 34-35)

These hearings paved the way for further dialogue on the subject of acquisition personnel management. Approximately nine months later, the House Investigations

Subcommittee held hearings to propose legislation for creating an acquisition corps that would undergo professional training, education, and career development. The Defense Acquisition Workforce Improvement Act was the result of these hearings.

C. DEFENSE ACQUISITION WORKFORCE IMPROVEMENT ACT

Managing defense acquisition programs requires competence in a broad spectrum of skills, including design, contract negotiation, cost estimating, auditing, production, testing and evaluation. Defense personnel are often at a distinct disadvantage when dealing with experienced government contractors. Fox (1988) noted that:

> . . .those assigned to key acquisition positions -- at most levels of DoD, from program managers to presidential appointees in the Pentagon -- are often unprepared for their jobs. These individuals lack the skills, training, and expertise required to manage the acquisition process effectively. (Cited in HEARINGS (i), 1990, p. 42)

The Defense Acquisition Workforce Improvement Act is specifically aimed at personnel management. The impact of people on a process is fundamental to its success. Former Rep. Nicholas Mavroules (D.-Mass.), House Investigations Subcommittee chairperson, emphasized this message. In his opening statement during the 1990 acquisition workforce Congressional hearings, he said:

> For many years, now, we have enacted all sorts of legislation dubbed 'Acquisition Reform.' We have changed the process. We have changed the procedures. We have changed many of the roles. But we have not yet addressed the most important element in the equation: people. (HEARINGS (ii), 1990, p. 1)

The subcommittee heard from a broad array of experts at the 1989 hearings on the acquisition workforce. Defense contractors, DoD executives, senior military officers, scholars, and members of Congress debated many workforce issues at the hearings before the

Act became law in November 1990. The five chapters that make up DAWIA are described below.

Chapter 1. General Authorities and Responsibilities: this chapter describes duties and responsibilities of key DoD acquisition executives. It requires the Secretary of Defense to set policies for career development. Additionally, this chapter established the important position of Director of Acquisition Education, Training, and Career Development in the Office of Secretary of Defense, and created a Director of Acquisition Career Management for each military department (DAWIA, 1990).

Chapter 2. Defense Acquisition Positions: this chapter requires the Secretary of Defense to designate all DoD jobs that are acquisition positions; it identifies acquisition career paths and mandates professional qualifications for acquisition personnel. According to ACT (1997), "The legislation would require cross fertilization of experience and talent among various acquisition career fields" (p. 7).

Chapter 3. Acquisition Corps: this chapter requires the Secretary of Defense to establish an Acquisition Corps[2]; it requires identification of "critical acquisition positions" for posts with significant responsibility; it sets a minimum service obligation for those positions; and sets education, training and experience requirements for critical acquisition positions. The previously-mentioned DAWIA summary noted that "the Acquisition Corps established as a result of this legislation would be a highly qualified cadre of individuals who, by demonstration of their capabilities . . . would have earned recognition as experts in the field of acquisition" (ACT, 1997, p. 3).

Chapter 4. Education and Training: this chapter establishes an internship, a cooperative education program and scholarships for entry-level personnel; it institutes a tuition reimbursement and training program for the acquisition workforce; and establishes a Defense Acquisition University to provide professional education and training. According to ACT (1997), "a major objective of the act is to increase the opportunities of civilian acquisition workforce personnel to compete for key, top-level management positions in the defense acquisition system" (p. 13).

Chapter 5. General Management Provisions: this chapter authorizes management information systems for each military department and defense agency; it requires annual workforce status reports for the Secretary of Defense; and it outlines reassignment policies, and minimum experience requirements (DAWIA, 1990).

---

[2]The Acquisition Corps consists of experienced acquisition professionals in the grades of Navy Lieutenant Commander, Army/Marine Corps Major, and civilian personnel in grades GS-13/GM-13 and above.

No where in the law are DAWIA's objectives explicitly stated. From the contents of the five chapters summarized above, DAWIA's objectives can be inferred to be:

- define acquisition career paths necessary for career progression
- establish professional qualifications
- provide workforce education and training
- increase the proportion of critical acquisition positions for civilians
- improve Program Manager proficiency
- reduce Program Manager turnover

The desired outcomes of DAWIA are to improve the way DoD acquires its goods and services, increase the professionalism of the acquisition workforce, and save taxpayer money. The Defense Acquisition Workforce Improvement Act is important because it "focus[es] on improving the effectiveness of the people who must implement the defense acquisition system and make it work" (ACT, 1997, p. 1).

D. WHY MEASURE PERFORMANCE?

Tracking and measuring performance is essential to achieving organizational objectives. Goal-setting is meaningless without follow-up to verify that the desired outcome was reached. Measurement is an inherent part of monitoring, controlling and improving a process or activity. Measurement must occur to manage and improve outcomes.

Some purposes for taking measures are:

- to know what is going on in the organization
- to identify performance disparities
- to provide feedback that compares performance to a standard

- to aid in making decisions regarding resources (Rummler & Brache, 1990).

The Training Resources and Data Exchange[3] (TRADE) Performance-Based Management Special Interest Group, listed similar reasons for measuring:

- Control: Measurements help to reduce variation.
- Self assessment: Measurements can be used to assess how well a process is doing, including improvements that have been made.
- Continuous improvement: Measurements can be used to identify defect sources, process trends, and defect prevention, and to determine process efficiency and effectiveness, as well as opportunities for improvement.
- Management assessment: Without measurements there is no way to be certain we are meeting value-added objectives or that we are being effective and efficient (TRADE, 1997, p. 3).

According to the Organisation for Economic Co-operation and Development[4] (OECD), an international public policy research institute based in Paris, "The main objective of performance measurement in public organisations is to support better management decision-making leading to improved outcomes for the community, and to meet external accountability requirements" (OECD, 1997, p. 1).

Another incentive for measuring performance in the public sector is the Government Performance and Results Act (GPRA). Passed by U.S. Congress in 1993, GPRA's objective is to improve the outcomes of federal programs. According to the Government Accounting

---

[3] The TRADE Performance-Based Management Special Interest Group was chartered by the U.S. Department of Energy to foster continuous improvement and facilitate the use of performance-based management techniques.

[4] OECD is based in Paris, France. The word "organisation," is spelled differently from common English spelling.

Office (GAO), GPRA "was enacted to improve the efficiency and effectiveness of federal programs by establishing a system to set goals for program performance and to measure results" (GAO (ii), 1997, p. 6). The Government Performance and Results Act requires federal agencies to develop strategic goals, measure performance against those goals, and report annually to Congress.

The next section summarizes performance measurement theory, measurement categories, and general measurement concerns.

E. PERFORMANCE MEASUREMENT THEORY

This section discusses important aspects of measurement that aid in focusing on what to measure, and provides suggestions to avoid measurement pitfalls. It begins with a definition of performance measures, and follows with a discussion of categories of measurement, measurement concerns, and measurement paradigms.

What are performance measures? "Performance measures quantitatively tell us something important about our products, services, and processes that produce them. . . . A performance measure is composed of a number and unit of measure. The number gives us a magnitude (how much) and the unit gives the number a meaning (what)" (TRADE, 1997, p. 2). Performance measures can be single-dimensional or multidimensional. Here is a simple example of a single-dimensional performance measure: $10. The number ten represents the magnitude, while dollars is the unit of measure. A multidimensional performance measure is expressed as a ratio of two or more units of measure. An example of a multidimensional performance measure is: 65 miles per hour.

1. Categories of Measurement

According to TRADE (1997), most performance measures can be grouped into six general categories or classifications: effectiveness, efficiency, quality, timeliness, productivity, and safety. The Organisation for Economic Co-operation and Development (1997) identified another performance measure category: financial results.

Each of the seven categories mentioned above measure performance, as compared to measuring compliance. Measures of compliance evaluate whether something was accomplished, while measures of performance evaluate how *well* something was accomplished.

Rhoads (1995) distinguished between measuring the implementation (i.e., compliance) of DAWIA and measuring the effectiveness of DAWIA. Measuring implementation is a much simpler process, and relatively straightforward. It is simply measuring compliance with the law. The basic question to be answered with regard to compliance is: have the statutory requirements of DAWIA been put in place? Specifically, have professional qualifications and career paths been established, as required by DAWIA?; have all acquisition positions been so designated, as required?; have critical acquisition positions been defined, and minimum service obligations been set?; have internships, co-ops, and scholarships been developed?; has Defense Acquisition University been established, and so forth. All of these are examples of measures that address implementation (or compliance) of DAWIA. Implementation metrics have also been referred to as "yes/no" metrics (TRADE, 1997) and "go/no go" metrics (Pope, 1997).

Measuring the effectiveness of DAWIA is a more difficult task than measuring implementation. According to Rhoads, this requires quantifying DAWIA's impact on the DoD acquisition process. In other words, what positive outcomes have taken place because of this law?

For example, chapter 3 of DAWIA mandates completion of a program management course before assignment as a Program Manager (PM). The implicit assumption is that a more highly trained PM is a more effective manager. To confirm this, some type of performance measure linked to a desired outcome has to be collected and analyzed. (The desired outcome in this example is effective program management.)

If a useful metric can be developed to measure this, then the effectiveness of this provision of DAWIA can be determined. This process can be applied by policy analysts to every DAWIA statutory requirement to find out if DAWIA has achieved its desired outcomes.

2. General Measurement Concerns

Measuring is a futile exercise if it is not done right, if the wrong thing is measured, or if no one uses the data. Basic measurement concepts are discussed in this section. Brinkerhoff and Dressler (1990) addressed four measurement concerns: validity, reliability, bias and sampling error. These concerns must be heeded to ensure quality measures. Two other measurement issues are judgmental versus nonjudgmental measures. Landy and Farr (1983) discussed the difference between these two types of measures, along with their strengths and weaknesses.

a. Validity

Validity is an important factor in measurement because it "determine[s] the 'fit' between the variable measured . . . and the construct about which we wish to make an inference". . . (Brinkerhoff & Dressler, 1990, p. 39). Restated in simpler terms, "Validity refers to the relationship between what is measured and what the person doing the measurement wants to know" (Brinkerhoff & Dressler, p. 38). Validity is established when the selected measure is related to an event or object that can be controlled or manipulated.

Brinkerhoff and Dressler explained the importance of validity by stating: "A measure of how much and how well a trainee performs on the job is a more valid measure of training's effectiveness than is an end-of-workshop opinion about training's quality" (p. 39). In this example, the metric is implied to be some sort of performance appraisal. The event that the investigator wants to know about is training quality. Clearly, an on-the-job measure of performance would be a more valid measure of training effectiveness than a survey of the trainee's opinions about the quality of training he or she received. An on-the-job performance measure would be related to the event (i.e., training effectiveness) that can be controlled, while an opinion survey would not.

b. Reliability

Reliability refers to consistency, accuracy and precision of a measure. Consistency is the measure's ability to produce the same results repeatedly. Accuracy involves the degree to which the measure reflects reality. The Training Resources and Data Exchange group defines accuracy as "The closeness of a measurement to the accepted true value. The smaller the difference between the measurement and the true value, the more

accurate the measurement" (TRADE, 1997, p. 13). Precision is defined as:

> The closeness of a group of repeated measurements to their mean value. The smaller the difference between the group of repeat measurements and the mean value, the more precise the instrument. Precision is an indicator of the repeatability, or consistency, of the measurement. (TRADE, p. 13)

Figure 2-1 depicts the relationship between precision and accuracy.

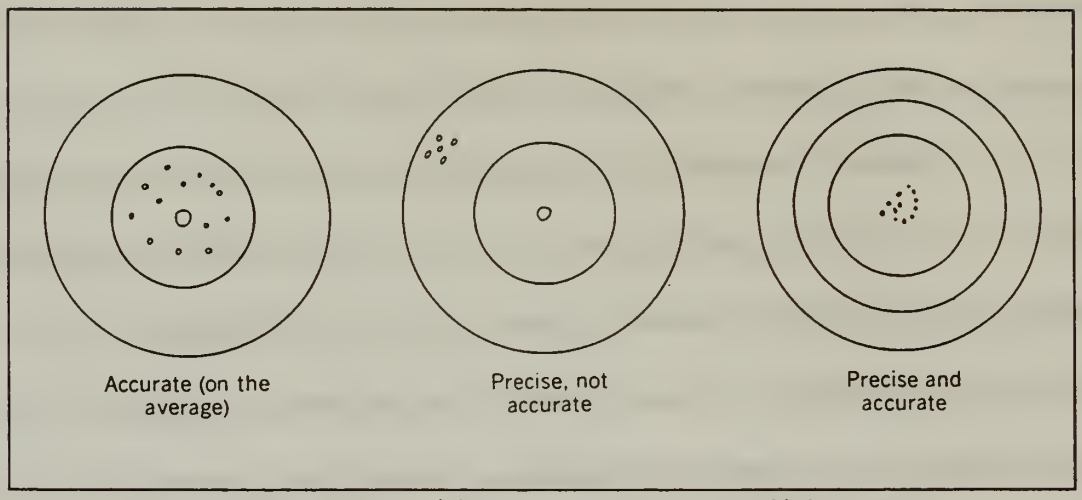

Figure 2-1. Precision vs. Accuracy. From Sink, p. 68.

c. Bias and Sampling Errors

Measures that compromise the outcome of what is being evaluated are biased. A bias in measurement is "A tendency or inclination of outlook that is a troublesome source of error in human sensing" (TRADE, p. 13). In this context, "sensing" refers to the measuring device (a person in this case) that can detect the presence of some phenomena.

Bias is a serious measurement concern that must be avoided. An example of bias is given here. A catalog merchandiser measures customer satisfaction by responses to customer questionnaires shipped with their products. If satisfied customers respond to the questionnaire at a higher rate than dissatisfied customers, the results may be biased.

Contented consumers may be more likely to fill out such a questionnaire, while dissatisfied consumers may simply shop elsewhere. If that situation was indeed true, then questionnaire results would tend to be favorably biased.

Sampling errors are flaws or misrepresentations of some type in a sample. According to Fremgen (1997), "The objective of sampling is to obtain evidence about the entire population, not just the sample" (p. 80). Sampling errors occur to some extent because, by definition, a sample comprises only a small portion of the population. Inevitably, the sample will not exactly represent the entire population, therefore some amount of error will always be present in any sample (Newbold, 1995). Generally, the larger the sample, the smaller the error.

d. Judgmental vs. Nonjudgmental Measures

Another measurement concern centers on the type of measure selected to evaluate performance. There are judgmental and nonjudgmental measures. Judgmental measures have been said to deal in abstractions, while nonjudgmental measures are based on tangible data. Judgmental measures are those that require discretion or judgment of the person taking a measurement. A personnel appraisal is an example of a judgmental measure because it requires subjective assessment of performance. Landy and Farr (1983) explained that "the rating process requires one individual to make a judgment about the performance level of another. This involves collecting information, weighing its value, and using it to make a statement about the performance of the person being rated" (p. 27). The Navy Fitness Report, for example, is a judgmental measuring device because it assesses the following (judgmental) measures: professional expertise, equal opportunity, military bearing, and

teamwork. Evaluating an individual using these measures requires discretion by the grader, and therefore, they are judgmental measures of performance.

Arguments against the use of judgmental measures assert that they fail to adequately measure performance. "Kane and Lawler (1979) contend that traits have no place in performance appraisal systems. They argue that traits (if we can measure them) are only characteristics of individuals that serve as causes or limiters of performance level and do not constitute performance per se" (Landy & Farr, 1983, pp. 86-87). Another significant weakness of judgmental measures is grader bias, both inadvertent and intentional. An example of a judgmental metric for DAWIA is a survey of the workforce regarding their knowledge of career paths and qualification requirements. Since a survey is prone to both researcher and respondent bias, this would not be an objective measure.

A nonjudgmental measure is definitive in nature. It is countable or quantifiable in some way, such as time, cost, quantity, weight, and so on. These types of measures are usually considered objective. Landy and Farr asserted: "The implication is that they are not open to interpretation, that they are unambiguous and reliable" (1983, p. 56). Whenever possible, nonjudgmental measures should be used to measure performance.

When developing and/or implementing a measurement system, the six measurement concerns just examined should not be ignored. The integrity of the resulting measurement data can be compromised if these measurement concerns are not observed.

3. Measurement Paradigms

It is also important to recognize organizational paradigms that hinder the ability to accurately track and measure performance. A paradigm is a customary pattern or model.

Manning and Ginzagr (1996) explained that "paradigms exist to give us rules and boundaries that help reduce uncertainty and make life predictable and more comfortable" (p. 11). An example of a paradigm is writing a document starting from the left side of the page and moving right. It could be done the opposite way (i.e., right to left), but it would be very confusing to most readers until that pattern became the accepted model. Some paradigms serve a useful function, such as the example just mentioned; some do not.

Sink and Tuttle (1989, pp. 58-60) identified several measurement paradigms that are dysfunctional in an organization. They are:

- *Measurement is threatening.* Workers often fear that new information about a process they perform or oversee may be used against them. If this were the case, workers might be inclined to undermine the measurement system.

- *Precision is essential to useful measurement.* Sink and Tuttle argue that performance measurement "does not have to be as precise as the measurement in a laboratory to be useful" (p. 59).

- *Single indicator focus.* An organization's overall "performance cannot be adequately explained or measured by a single indicator" (p. 59). When single indicators are used, there is a tendency to overreact to measurement results.

- *Subjective measures are sloppy.* Landy and Farr's (1983) discussion against the use of judgmental measures supports this paradigm. In contrast, Sink and Tuttle argued that subjective measures *can* be reliable. Their first argument is that industrial psychologists have refined measurement technology for measuring so-called "soft" performance dimensions to some degree of accuracy. Their second argument in favor of subjective measures is bolstered by the common practice of distributing bonuses and rewards only to divisions that show tangible results; this practice minimizes the contributions of support divisions, and inhibits teamwork, cooperation, and gainsharing.

- *Standards operate as ceilings on performance.* Standards inherently imply performance that meets a desired level is good enough, therefore they remove incentives for continual improvement.

These paradigms can be damaging to the measurement process, and can work against improving results. Sink and Tuttle (1989) noted that "some of these paradigms get in the way of the organization's ability to continuously improve performance" (p. 58). These and other detrimental measurement paradigms must be recognized and eliminated to ensure a measurement system works as designed. A good performance measurement approach will consider each measurement concern and paradigm, to ensure that measures of effectiveness are valid, accurate, reliable and unbiased.

F. METRICS

This section defines metrics and discusses attributes of useful metrics. General guidelines are given on how to develop a metric. Figure 2-2 depicts classification of metrics using five of the seven performance categories previously discussed in the performance measurement theory section. And finally, examples of DAWIA metrics are shown in Figure 2-3.

1. Definition of Metrics

Researchers use various terms to refer to the concept of measure. Rhoads (1995) used "measures of effectiveness," while Chang and De Young (1995) employed the phrase "key indicator." Landy and Farr (1983), Brinkerhoff and Dressler (1990), and Sink and Tuttle (1989) all preferred the term "measure." The Training Resources and Data Exchange group distinguished between measures, which they defined as "quantitative evaluations of the products or services of a process or system," and metrics which "are standards of measurement (such as length, area, frequency, mass, and so on)" (TRADE, 1997, p. 15). The military's buzz word for measure is metric. The terms measure, metric, measure of

effectiveness and key indicator are used interchangeably throughout this study.

Webster's Dictionary (1994) defines metrics as "the branch of prosody dealing with measure and metrical structures." Prosody is the study of verification, according to Webster. This definition essentially means that metrics involve a system of verification using measures.

The Metrics Handbook (METRICS) by U.S. Air Force Systems Command provides a more user-friendly description of metrics. It states:

> Metrics are nothing more than meaningful measures. For a measure to be meaningful, however, it must present data that allow us to take action. . . . Metrics foster process understanding and motivate action to continually improve the way we do business. This is distinguished from measurement, in that, measurement does not necessarily result in process improvement. Good metrics always will. (1991, p. 1-1)

2. Attributes of Metrics

According to METRICS (1991), to be a useful measure of effectiveness, a metric should:

- be meaningful to the customer
- relate to organizational goals
- be simple, understandable, logical and repeatable
- show a trend
- be clearly defined
- be economical to collect
- be timely
- drive the appropriate action.

Chang and De Young (1995) included the following additional criteria for effective metrics. Their recommendations for metrics are as follows. Metrics should:

- contain both effectiveness and efficiency components
- focus on success rather than failure
- focus on continuous improvement, rather than compliance only.

The Training Resources and Data Exchange group (1997) identified two other characteristics: a metric must provide a basis for decision making, and it must be compatible with existing measurement systems. Many metric attributes mentioned in this section will be used in the metric template introduced in Chapter IV.

The remainder of this section deals with metric development -- how to come up with effective measures, followed by a discussion of metric classification, and finally examples of metrics.

3. Metric Development

Developing metrics is instrumental to activating a performance measurement system. According to TRADE (1997), "Performance metrics should be constructed to encourage performance improvement, effectiveness, efficiency, and appropriate levels of internal controls" (p. 23). Pinker, Smith and Booher (1997) suggested these five guidelines for coming up with new metrics:

- Identify what is to be evaluated (i.e., what process or activity will be measured?)
- Identify relevant properties of what is to be measured (i.e., quality, effectiveness, efficiency, etc.)
- Identify types of potential metrics (list will be pared down if metrics are found to be undesirable -- too costly, not timely, unrelated to organizational goals, etc.)
- Select metrics and be able to justify those selections.

- Put boundaries on what is being measured.

The Training Resources and Data Exchange group presented a slightly different, but effective approach to developing metrics. The steps are listed below:

- decide the outcomes wanted
- describe the major work processes involved
- identify the key results needed
- establish performance goals [i.e., standards] for the results
- define measures for the goals
- select appropriate measures (TRADE, 1997, p. 16)

4. Classification of Metrics

Figure 2-2 shows how TRADE classified performance metrics using five of their six categories of measurement, which were discussed earlier in this chapter. The five categories

Measure of...	Measures...	Expressed as ratio of...
Efficiency	Ability of an organization to perform a task	Actual input/ planned input
Effectiveness	Ability of an organization to plan for output from its processes	Actual output/planned output
Quality	Whether a unit of work was done correctly. Criteria to define "correctness: are established by the customer(s).	Number of units produced correctly/total number of units produced
Timeliness	Whether a unit of work was done on time. Criteria to define "on-time" are established by the customer(s).	Number of units produced on time/total number of units produced
Productivity	The amount of a resource used to produce a unit of work	Outputs/inputs

Figure 2-2. Classification of Performance Metrics. From TRADE, p. 24.

shown in the first column are effectiveness, efficiency, quality, timeliness, and productivity. The second column describes what is being measured, and the third column gives an example of a metric. The data in Figure 2-2 will assist the user in developing appropriate metrics according to what type of measurement data is being sought.

5. Examples of Metrics

Now that you know what a metric is, how to develop one, and what attributes are important in selecting effective metrics, some metric examples will be given in this section. Some examples of Pope's (1997) metrics are shown in Figure 2-3. Pope developed metrics (shown in second column) for measuring activities needed to meet a given DAWIA objective

DAWIA Objective	Metric	Metric Category
provide career paths for acquisition workforce	1. have acquisition positions been designated? 2. percent of personnel in intern program 3. percent of personnel in scholarship program	implementation effectiveness effectiveness
establish professional qualifications	1. percent of workforce meeting certification requirements	implementation
provide workforce education & training	1. has a scholarship program been established? 2. number of training & education opportunities available since DAWIA 3. amount of training received since DAWIA	implementation implementation implementation
increase proportion of critical acquisition positions for civilians	1. percent of senior-level positions that civilians are permitted to occupy since DAWIA 2. percent of senior-level positions occupied by civilians	implementation implementation
improve Program Manager (PM) proficiency	1. average amount of acquisition experience PMs possess 2. promotion rates of military PMs compared to military promotion rates as a whole	effectiveness effectiveness
reduce PM turnover	1. average duration a Program Manager serves	effectiveness

Figure 2-3. Suggested DAWIA Metrics. From Pope, pp. 60-66.

(shown in first column). The third column in Figure 2-3, labeled "Metric Category," was incorporated to Pope's work; it indicates metric categories (i.e., performance metrics vs. implementation metrics).

G. CHAPTER SUMMARY

Although several defense acquisition studies and commission reports over several decades recognized the need for training and educating the acquisition workforce, their findings were not implemented until enactment of DAWIA. The Defense Acquisition Workforce Improvement Act is a federal law aimed at improving the quality of the defense acquisition workforce. It attests to the fact that people are the key element to success in procurement. The main objective of DAWIA is development of the workforce through training, education, and professional certification.

Measuring performance is essential to achieving organizational objectives. Unless measures are taken, there is no assurance that the desired outcome was reached. A performance measure is a quantitative assessment of some phenomenon of interest. It has two components: a number and unit of measure. The number indicates magnitude and the unit gives the number a meaning. Measures typically fall into one of eight measurement categories: effectiveness, efficiency, quality, timeliness, productivity, safety, financial results, and compliance.

For measurement to be useful, measures must be valid, reliable, and unbiased. Measures are either judgmental and nonjudgmental. Judgmental measures, such as such as a personnel appraisal, are those that require discretion by the person taking the measurement. A nonjudgmental measure is more objective in nature, such as time, cost, quantity, or weight.

To measure the impact of DAWIA, metrics must be employed. A metric is nothing more than a meaningful measure that provides useful information about a process or activity. Effective metrics have attributes that add value to the measurement in some way. Examples of metric attributes are: meaningful, simple, understandable, and decision-oriented.

In developing metrics, a good way to start is to identify and place boundaries on what is to be measured. According to TRADE, if you can not describe what is to be measured, you can not improve it.

The next chapter will introduce four performance measurement approaches, and discuss strengths and weaknesses of each. An analysis of the four approaches will be conducted to determine which approach is most suitable for measuring DAWIA.

III. PERFORMANCE MEASUREMENT APPROACHES

A. INTRODUCTION

A performance measurement approach is used in this study to identify metrics that are accurate and/or effective indicators of acquisition workforce performance. The process or approach will link metrics with specific performance outcomes. Several important modeling principles that can be applied to the performance measurement process are also discussed. Four performance measurement approaches are then presented along with a discussion of their strengths and weaknesses. Finally, a measurement approach is selected and defended.

A performance measurement approach is a step-by-step method for transforming measurement data into useful and timely information that can be used to improve a process or activity. In contrast, a model is "a representation of reality or of a real-life situation" (Render & Stair, 1997, p. 14). The approaches presented in this chapter are not models per se, though they are sometimes loosely referred to as such. They are more accurately described as a framework, or simply methods for confronting a task. The terms approach, method, process and framework are used interchangeably in referring to any of the four approaches. Only with the Measurement Linkage Model (Chang & De Young, 1995) is the term "model" used, as that is how the authors refer it.

Several basic modeling principles can be applied to the performance measurement approaches. According to Liao (n.d.), the following principles are useful in modeling: Simplification, capturing the essence of the system, relevance, and accuracy.

Simplification is essential to making a complex system understandable and manageable. According to Wallace (1994), an important principle in model-building is that it should be understandable. "Its underlying assumptions must be clearly stated, so that the user can verify its appropriateness to the situation at hand" (p. 232). A model should be as simple as possible without compromising its essence. This leads us to the second concept. To be a true representation of reality, the defining characteristics of the situation or system must be preserved. This is what Liao (n.d.) described as capturing the essence of the system.

Relevance is another key modeling principle. Liao (n.d.) noted that "The model should only include those aspects of the system [or situation] that are relevant to the study objectives. Other aspects should be simplified as much as possible or assumed away if possible" (p. 4). The last principle is accuracy. "The degree of accuracy of the information gathered for the model should be considered according to its relevance to the decision" (Liao, p. 4).

Based on the modeling principles cited above, the approach should be simple, accurate and understandable; it should be relevant to the objectives of the study; and it must preserve the essence of the situation it represents. Each of these guidelines can be applied to evaluate the performance measurement approaches that will be introduced in the next section.

B. ALTERNATIVE APPROACHES

Four approaches are presented and examined below, followed by a discussion of their strengths and weaknesses. The four approaches are: The General Measurement Methodology by Sink and Tuttle (1989), the TRADE (1997) Performance Measurement Process, Chang and De Young's (1995) Measurement Linkage Model, and the Metric Development Process

of the U.S. Air Force Systems Command (1991).

1. General Measurement Methodology

Sink and Tuttle (1989) advocated the basic flow process shown in Figure 3-1. They considered their methodology a tool for converting measurement data into measurement information. It is one of their various techniques to "collect, store, process, retrieve, and portray data about the performance of various types of organizational systems . . . " (p. 251). The General Measurement Methodology (Sink & Tuttle, 1989, p. 254-257) consists of five phases, some involving multiple steps.

Phase 0. Preparation.
Step 1. Form a measurement development team. The team will design and carry out the measurement system.
Step 2. Create a suitable climate for measurement.

Phase 1. What to measure.
Step 1. Develop performance measures.
Step 2. Audit the measures. The purpose of auditing is to ensure quality and usefulness of metrics.
Step 3. Break down the measures. The purpose of breaking them down into attributes is to identify something that is countable.

Phase 2. Develop the measurement process.
Step 1. Select measurement technique(s) such as statistical process control, cost/benefit analysis, time and motion study, etc.
Step 2. Form a measurement design team to develop a way to apply the technique.
Step 3. Ensure the technique can be carried out.
Step 4. Adapt the measurement techniques to fit the situation, if necessary.

Phase 3. Collect data.
Step 1. Identify data sources.
Step 2. Eliminate infeasible measures that may be too costly or too difficult to collect.
Step 3. Decide data collection, storage and retrieval methods.
Step 4. Identify personnel to collect and record data.
Step 5. Collect data.

Phase 4. Validate process output. Sink and Tuttle view this phase as the "shakedown period." Discover how to use the new information; recalibrate the measurement system if necessary; and figure out how to link the results to action and decisions.

Phase 5. Link to improvement. "The key to success is to link measurement to improvement" (p. 256). Two examples of how to link measurement to improvement are: Use of control charts that show when to take action, and measures linked to a reward system.

Figure 3-1 shows the General Measurement Methodology. This method provides an understandable and comprehensive approach to performance measurement.

One strength of the General Measurement Methodology is the design team concept. At the beginning, a team is formed to develop and carry out the measurement system. It is the team's responsibility to create a suitable organizational climate for measurement. Sink and Tuttle (1989) acknowledged that this is not easy. "Creating a climate that will support measurement for improvement is a critical step in our general methodology, and will play a big role in the success or failure of your attempts to build improved measurement systems . . ." (p. 254). The practice of involving several people in developing a measurement system lends credibility to the process, and creates a "buy-in" atmosphere. Since the team becomes stakeholders in the process, they will have an investment in its outcome. Sink and Tuttle (1989) found that measurement teams "ended up with high-quality measurement systems, greater acceptance of the results, and a better foundation for moving through to implementation of the measurement systems" (p. 232).

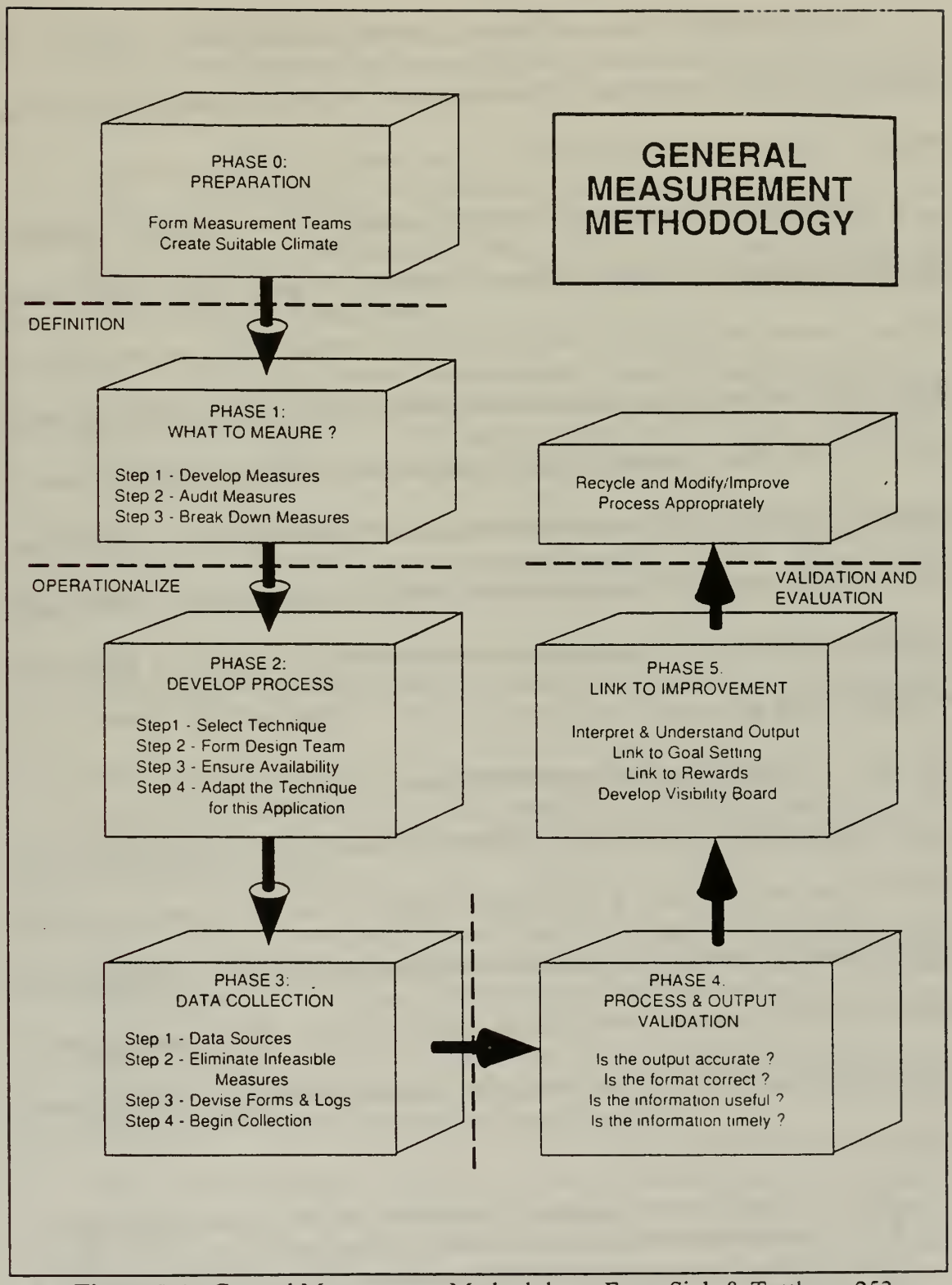

Figure 3-1. General Measurement Methodology. From Sink & Tuttle, p. 253.

Another strength of Sink and Tuttle's approach is the audit process in phase 1. The ultimate objective in using a measurement system is to acquire good measurement data that provides a basis for decision-making; this is what the auditing process provides.

A weakness of Sink and Tuttle's approach is that it does not compare measurement data against performance standards. Sink and Tuttle argued against the use of performance standards because they do not fit in with the concept of continuous improvement. They said that a "standard carries the connotation of desired level. If the standard is met, performance is OK" (p. 60). Other literature stresses the importance of having standards by which to compare actual performance to expected performance. The Training Resources and Data Exchange group noted that "having goals and standards is the only way to meaningfully interpret the results of your measurements and gauge the success of your management systems" (1997, p. 4).

2. Performance Measurement Process

The Training Resources and Data Exchange Performance Measurement Process is the result of a group study by the U.S. Department of Energy. The study produced "reference material to assist in the development, utilization, evaluation, and interpretation of performance measurement techniques and tools to support the efficient and effective management of operations" (TRADE, 1997, n.p.). Its 11 steps are described below, along with a diagram of the process, shown in Figure 3-2.

> *Step 1. Identify the process flow.* This is the first and perhaps most important step. If your employees cannot agree on their processes, how can they effectively measure them or utilize the output of what they have measured? (p. 4). Output of step 1: A list of key processes and flow diagrams for each.

Step 2. Identify the critical activity to be measured. Critical activities are those that significantly impact total process efficiency, effectiveness, quality, timeliness, productivity, or safety (p. 5). Output of step 2: A list of critical activity areas for the key process.

Step 3. Establish performance goals or standards. All performance measures should be tied to a predefined goal or standard. . . . Having goals and standards is the only way to meaningfully interpret the results of your measurements and gauge the success of your management systems (p. 4). Output of step 3: A list of goals for each critical activity within the process.

Step 4. Establish performance measurements. Translate what you want to know into a performance measure; identify the raw data that will generate the performance measure; identify location of data, and who will gather it; decide frequency of measurements. Output of step 4: Performance measures.

Step 5. Identify responsible party(s). Decide who will collect, analyze and report data, who will determine if corrective action is needed, and who will make those changes. Output of step 5: A list of people and their areas of responsibility.

Step 6. Collect data. Besides writing down the numbers, the data need to be pre-analyzed in a timely fashion to observe any early trends and confirm the adequacy of your data collection system. Output of step 6: Data collection forms and system.

Step 7. Analyze/report actual performance. In this step, the raw data are formally converted into performance measures, displayed in an understandable form, and disseminated in the form of a report (p. 4). Output of step 7: A presentation of the data in the form of a report.

Step 8. Compare actual performance to goals. In this step, compare performance, as presented in the report, to predetermined goals or standards and determine the variation (if any) (p. 4). Output of step 8: Decision based on performance variance.

Step 9. Determine if corrective actions are necessary. Step 9 is a decision step. You can either change the process or change the goal (p. 12). Output of step 9: An action plan to implement changes or reevaluate goals.

Step 10. Make changes to bring process back in line with goal. This step only occurs if corrective action is expected to be necessary. The actual determination of the corrective action is part of the quality improvement process, not the performance measurement process. This step is primarily concerned with improvement of your management system (p. 4). Output of step 10: A successfully implemented plan.

Step 11. Determine if new goals or measures are needed. Goals need to be challenging, but also realistically achievable. Changes in performance measures and goals should be considered annually and integrated into planning and budgeting activities (p. 12). Output of step 11: New goals, measures, or no change.

The TRADE performance measurement method is arguably the most comprehensive of the four methods. Its strongest attribute is a specified output at each step in the process. This leaves no doubt about what is supposed to happen during each step. Also, the TRADE method is the only one to flowchart the process that is to be measured. The TRADE group explained that "This is the first and perhaps most important step. If your employees cannot agree on their process(es), how can they effectively measure them or utilize the output of what they have measured?" (1997, p. 4). Another reason for flowcharting the process is "individuals will receive a new understanding of their processes. As participants, you can count on their later support to make the performance measurement system work" (p. 5).

Another important advantage of this method is step 5, which requires specific identification of persons responsible for specific actions. This clarifies roles and relationships, ensures actions are carried out as expected, and avoids finger-pointing. In short, it leaves no one "in the dark."

3. Measurement Linkage Model™

The Measurement Linkage Model, is a trademark of Richard Chang Associates (Chang & De Young, 1995, p. 16). They designed their model for use by a division or work group within an organization.

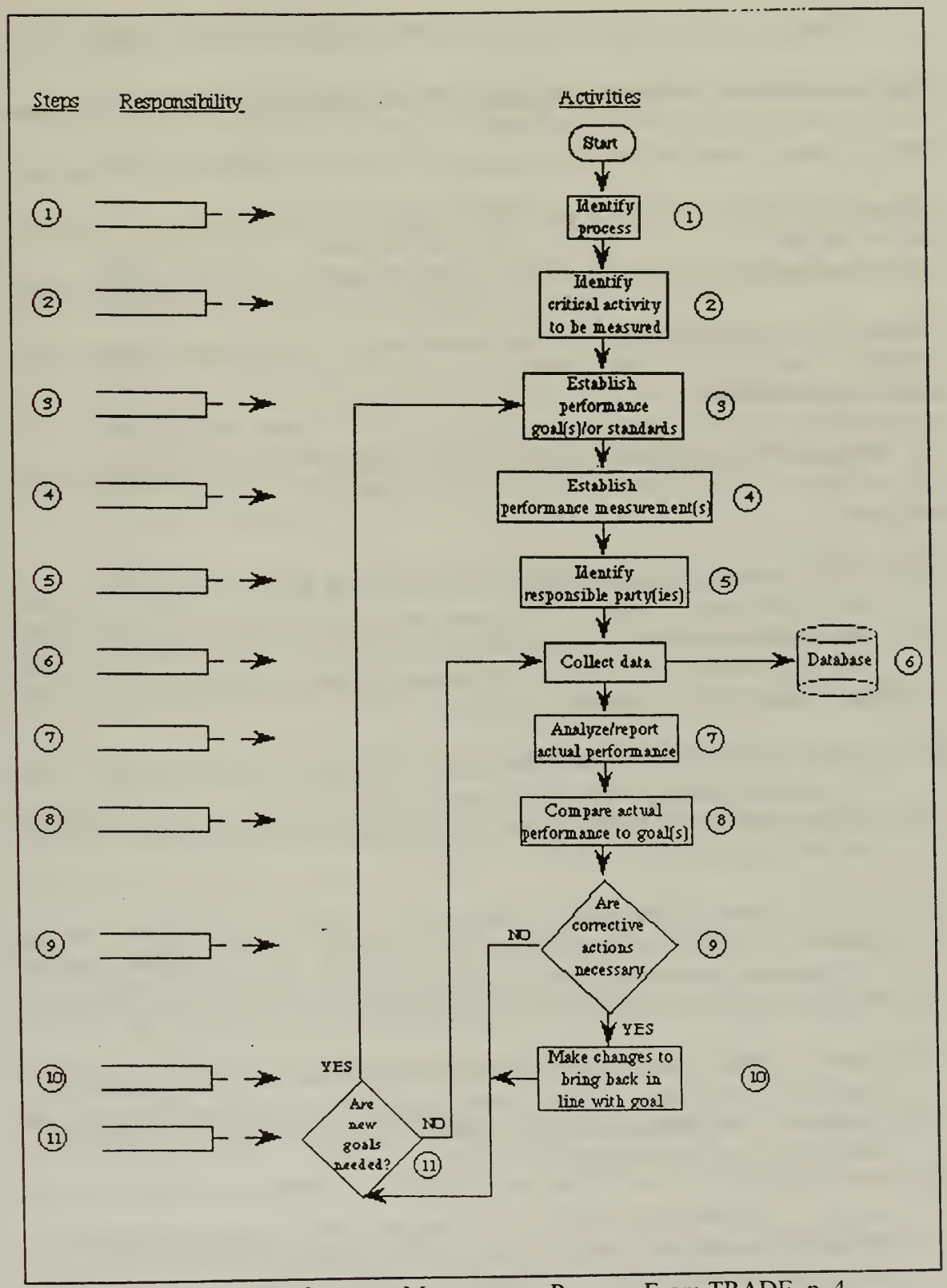

Figure 3-2. Performance Measurement Process. From TRADE, p. 4.

Chang and De Young used the terms "key result area," "key indicator," and "performance target," which are fundamental to understanding their "model." A key result area is a "make or break" performance category that is crucial to organizational success, such as customer loyalty. "A Key Indicator (KI) is a specific measure which helps to determine how well you are performing in a given KRA [key result area]" (p. 17). An example of a key indicator of customer loyalty is customer satisfaction ratings. A performance target is simply a performance standard.

Figure 3-3 depicts the Measurement Linkage Model™. Here is a description of its eight steps.

> *Step 1. Develop organization-wide KRAs [key result areas], KIs [key indicators], and performance targets.* An organization's vision, mission, and strategic plan provide direction for work groups throughout the organization. An executive team must first develop KRAs, KIs and performance targets to quantify its strategic plan into measurable components . . . . (p. 17)
>
> *Step 2. Select organization-wide KRAs and KIs linked to your work group.* All work groups produce products and/or services in support of the organization's goals. Work groups [must] select the appropriate organization-wide KRAs and KIs to which they directly contribute. (p. 19)
>
> *Step 3. Develop work group key result areas.* . . . [Work groups] examine "the big picture" of their work group as a system of suppliers, inputs, processes, outputs, customers and customer requirements. Good KRAs answer the question, "Which are the most critical work group performance outcomes we must achieve?" (p. 19)
>
> *Step 4. Develop work group key indicators.* KIs break each key result area into measurable components. KIs are the "yardsticks" by which one can measure progress. . . . Good KIs answer the question, "How will a work group know if it is making progress toward its KRAs?" (p. 20)
>
> *Step 5. Determine data collection, tracking and feedback methods.* Methods for KI data collection and feedback are critical in determining the success of the work group's measurement system. This step determines how, where, and when data will be collected. Determine who is accountable for monitoring, reporting, and using the

results to make decisions. (p. 20)

Step 6. Gather baseline data and set performance targets. Baseline data provides an excellent source of information to help identify [realistic] performance levels (or targets). (p. 20)

Step 7. Establish work group objectives and tactics. Work groups identify improvement opportunities and make plans for action. This is the "road map" to help achieve the targets. Once a work group makes improvement plans, they need to take specific action to improve performance. (p. 21)

Step 8. Implement plans, monitor performance and provide feedback. Continually monitor performance of the work group. Provide feedback to appropriate managers/employees so they can continue to improve the work group. Monitor the measurement system to ensure the work group is continuing to measure the right things. Adjust the measurement system as the business shifts its focus and responds to an ever-changing world. (p. 21)

One advantage of the Measurement Linkage Model™ is its flexibility. The authors intended for this model to be adapted to specific organizational needs. Their model allows for changes in the external and internal environment, such as customer needs, regulations, competition, etc. These changes frequently affect an organization's priorities and focus, which in turn, would likely influence key result areas and key indicators. This feature is represented in the diagram's top right corner.

A second important strength is its linkage between key indicators (metrics) and key result areas. This connection ensures that what is important to the organization is what gets measured. Another notable feature is its continuous loop from step 8 back to steps 1 and 2. This results in a continuous review of measurement results against performance categories that are crucial to organizational success (i.e., key result areas). The loopback is important because it focuses attention to ensure that the right things get measured. Another loop cycles back from step 8 to step 7. This results in continuous improvement by directing the user to

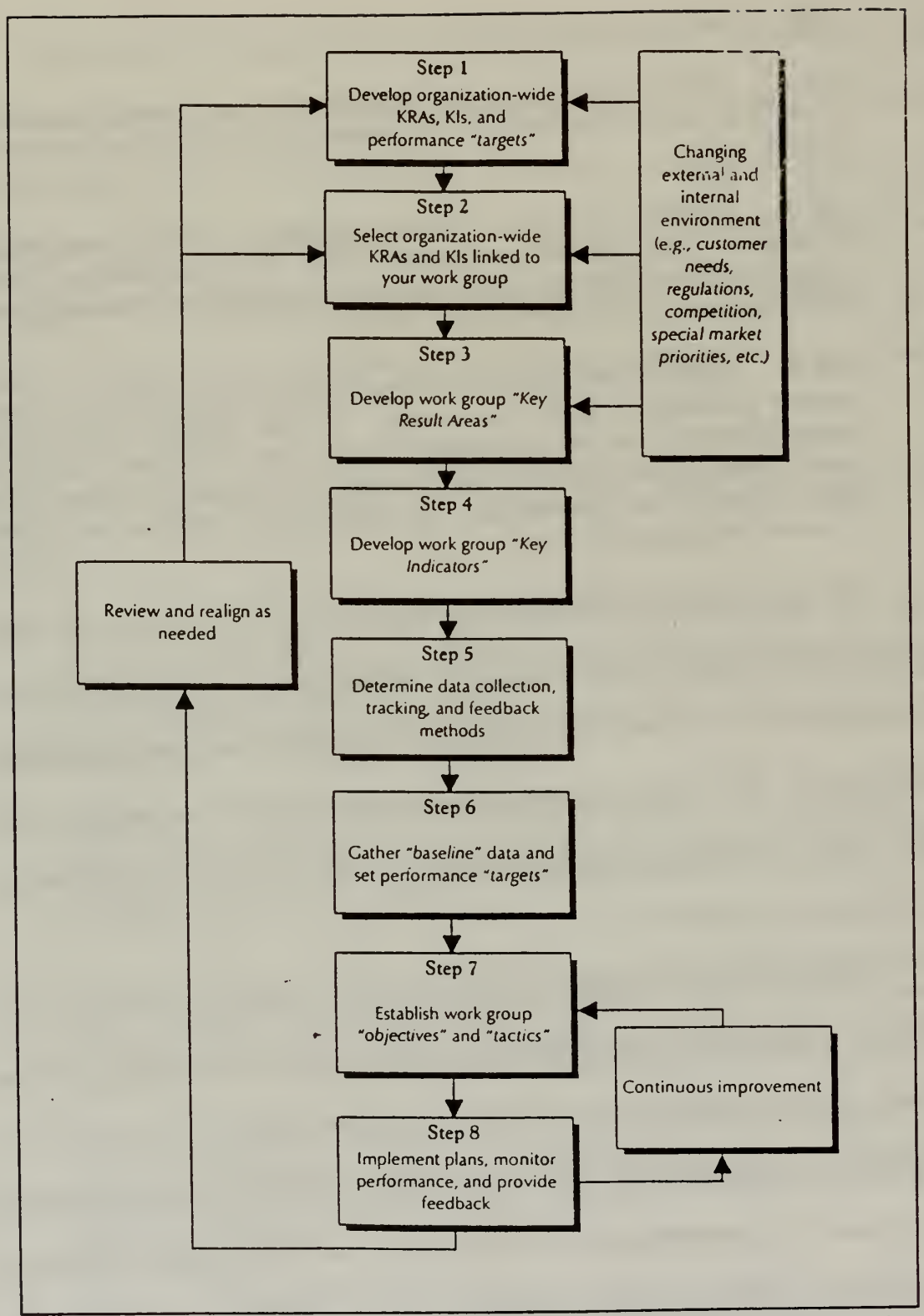
Step 1
Develop organization-wide KRAs, KIs, and performance "targets"
Step 2
Select organization-wide KRAs and KIs linked to your work group
Step 3
Develop work group "Key Result Areas"
Step 4
Develop work group "Key Indicators"
Step 5
Determine data collection, tracking, and feedback methods
Step 6
Gather "baseline" data and set performance "targets"
Step 7
Establish work group "objectives" and "tactics"
Step 8
Implement plans, monitor performance, and provide feedback
Changing external and internal environment (e.g., customer needs, regulations, competition, special market priorities, etc.)
Review and realign as needed
Continuous improvement

Figure 3-3. Measurement Linkage Model™. From Chang & De Young, p. 16.

continually identify new opportunities to improve after seeing the results of previous action.

4. Metric Development Process

Developed by the U.S. Air Force Systems Command, the Metric Development Process (METRICS, 1991, pp. 3-3 - 3-5) combines metric development with a performance measurement system. Figure 3-4 shows a diagram of the Metric Development Process, along with a practical example of how it is used. It entails the following ten steps.

> *Step 1. Identify your purpose.* It is important to first align your purpose with your organization's mission, vision, goals, and objectives. These should be inextricably linked to meeting customer needs and serve as a foundation for accomplishing and sustaining continuous, measurable improvement.
>
> *Step 2. Develop an operational definition starting with your customer.* Define the who, what, when, why and how of this metric in sufficient detail to permit consistent, repeatable and valid measurement to take place. . . . (p. 3-3)
>
> *Step 3. Identify and examine existing measurement systems.* Once the link to objectives and goals has been established, it is essential to determine if existing metrics or other measurement systems exist that satisfy your requirements. Don't "reinvent the wheel." Use existing process measurements when they exist. (p. 3-3)
>
> *Step 4. Generate new metrics if existing metrics are inadequate.* . . . We are interested in those upstream process measures which drive the final outcome and are the key to making process improvements. The assumption is: if process performance is monitored and improved, the quality of the products and services will improve. (p. 3-3)
>
> *Step 5. Rate your metric against the "eight attributes" of a good metric.* . . . If you feel your metric sufficiently satisfies these criteria for a good metric, go on to Step 6. If not, return to Step 2 and correct the deficiencies. (p. 3-4)
>
> *Step 6. Select appropriate measurement tools.* Select the proper tool for analyzing and displaying your data. [Common statistical process control tools include control charts, cause and effect diagrams, pareto charts, and histograms]. . . . (p. 3-4)
>
> *Step 7. Baseline your process.* Start acquiring metric data. This serves as a baseline for determining the capability of your process. . . . (p. 3-4)

Step 8. Collect and analyze metrics data over time. Continue aggregating metric data over time. Examine trends. . . . (p. 3-4)

Step 9. Finalize the metric presentation. Based on the results of the previous steps, you are finally ready to present the metric externally. . . . The appropriate level of detail should be determined by discussion with the customer. . . . (p. 3-4)

Step 10. Initiate process improvement activities. . . . Once improvements have been implemented, the process above may start over or it may pick up again at almost any step. Remember, continuous improvement requires continuous effort. This step in the development process is the most critical for your improvement efforts to become a reality. . . . (p. 3-5)

In conjunction with the Metric Development Process just described, METRICS introduced the "metric package." A metric package consists of three things: "the operational definition [of the metric], the actual measurement and recording of data, and the metric presentation" (p. 3-1). These three elements are described in detail below, and are depicted in Figure 3-5.

The operational definition is the first element of a metric package. An operational definition, according to METRICS "is the precise explanation of the process being measured" (p. 3-1). Deming, the late quality management guru, explained that an operational definition is needed to clearly communicate an ambiguous concept. He explained:

> Adjectives like good, reliable, uniform, round, tired, safe, unsafe, unemployed have no communicable meaning until they are expressed in operational terms of sampling, test, and criterion. The concept of a definition is ineffable: It cannot be communicated to someone else. . . . An operational definition is one that people can do business with. (Deming, 1982, pp. 276-277)

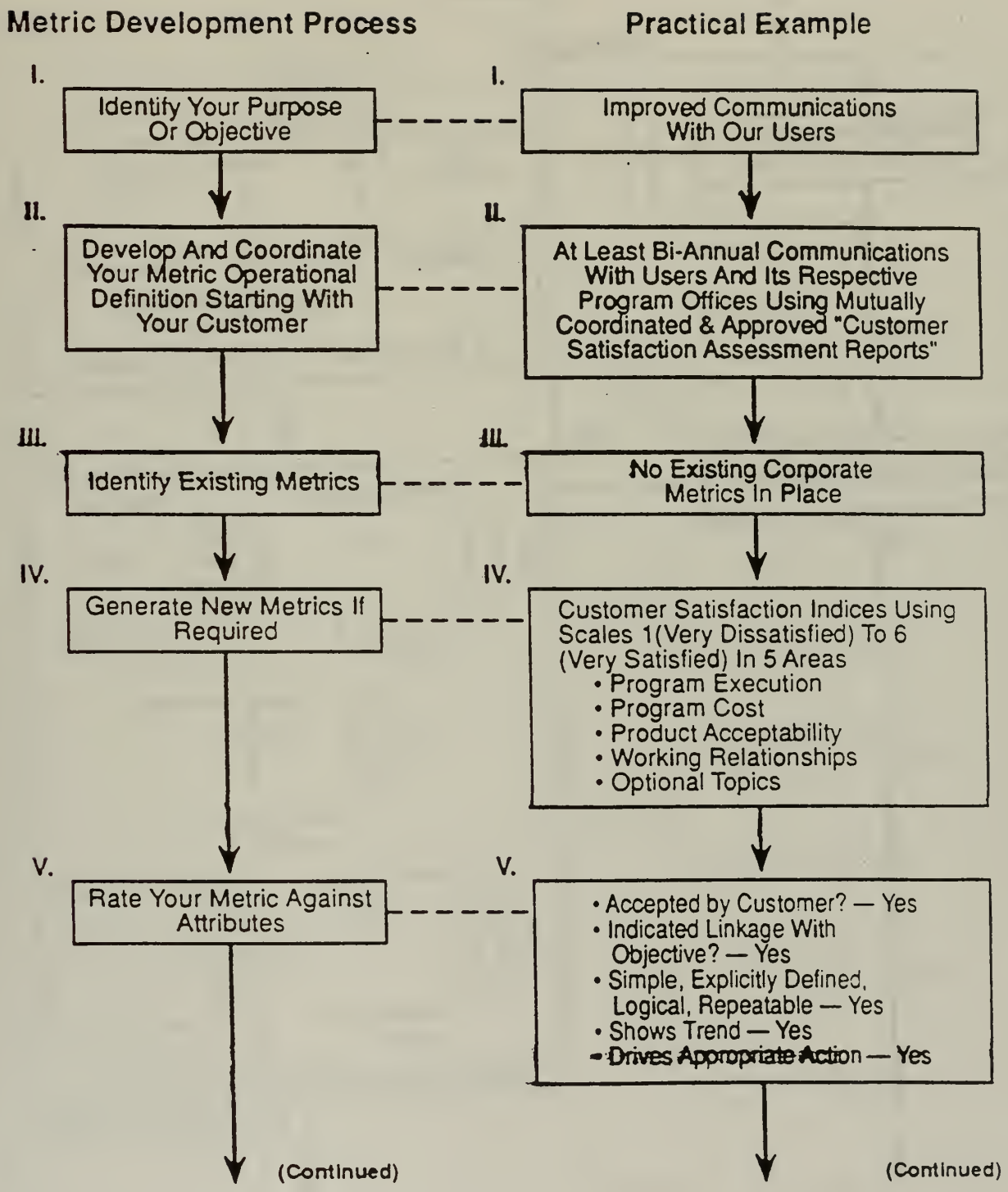
Metric Development Process
Practical Example
I.
Identify Your Purpose Or Objective
I.
Improved Communications With Our Users
II.
Develop And Coordinate Your Metric Operational Definition Starting With Your Customer
II.
At Least Bi-Annual Communications With Users And Its Respective Program Offices Using Mutually Coordinated & Approved "Customer Satisfaction Assessment Reports"
III.
Identify Existing Metrics
III.
No Existing Corporate Metrics In Place
IV.
Generate New Metrics If Required
IV.
Customer Satisfaction Indices Using Scales 1(Very Dissatisfied) To 6 (Very Satisfied) In 5 Areas
• Program Execution
• Program Cost
• Product Acceptability
• Working Relationships
• Optional Topics
V.
Rate Your Metric Against Attributes
V.
• Accepted by Customer? — Yes
• Indicated Linkage With Objective? — Yes
• Simple, Explicitly Defined, Logical, Repeatable — Yes
• Shows Trend — Yes
• Drives Appropriate Action — Yes
(Continued)
(Continued)

Figure 3-4. Metric Development Process. From METRICS, p. 3-6.

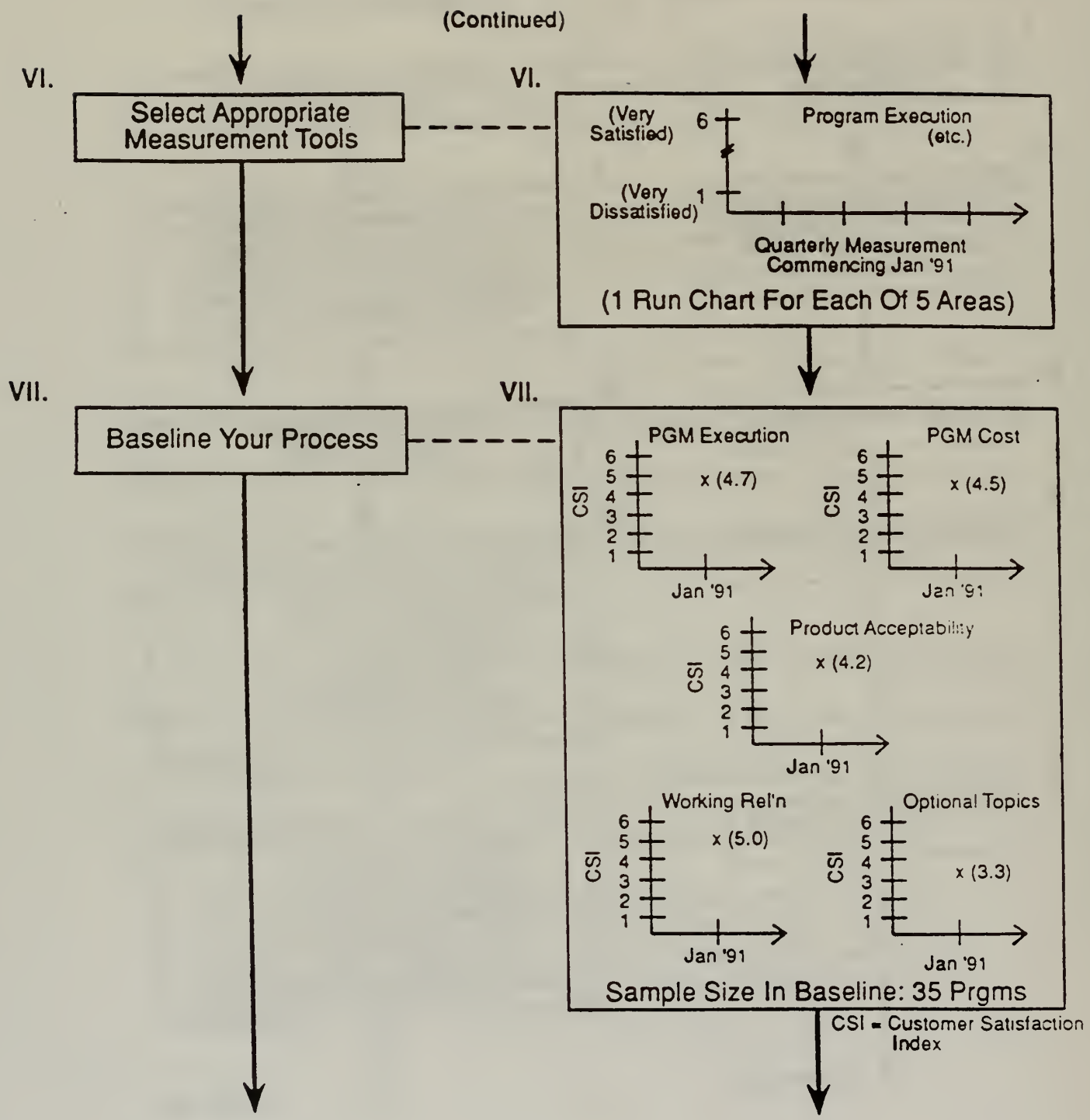

Figure 3-4 (cont). Metric Development Process. From METRICS, p. 3-7.

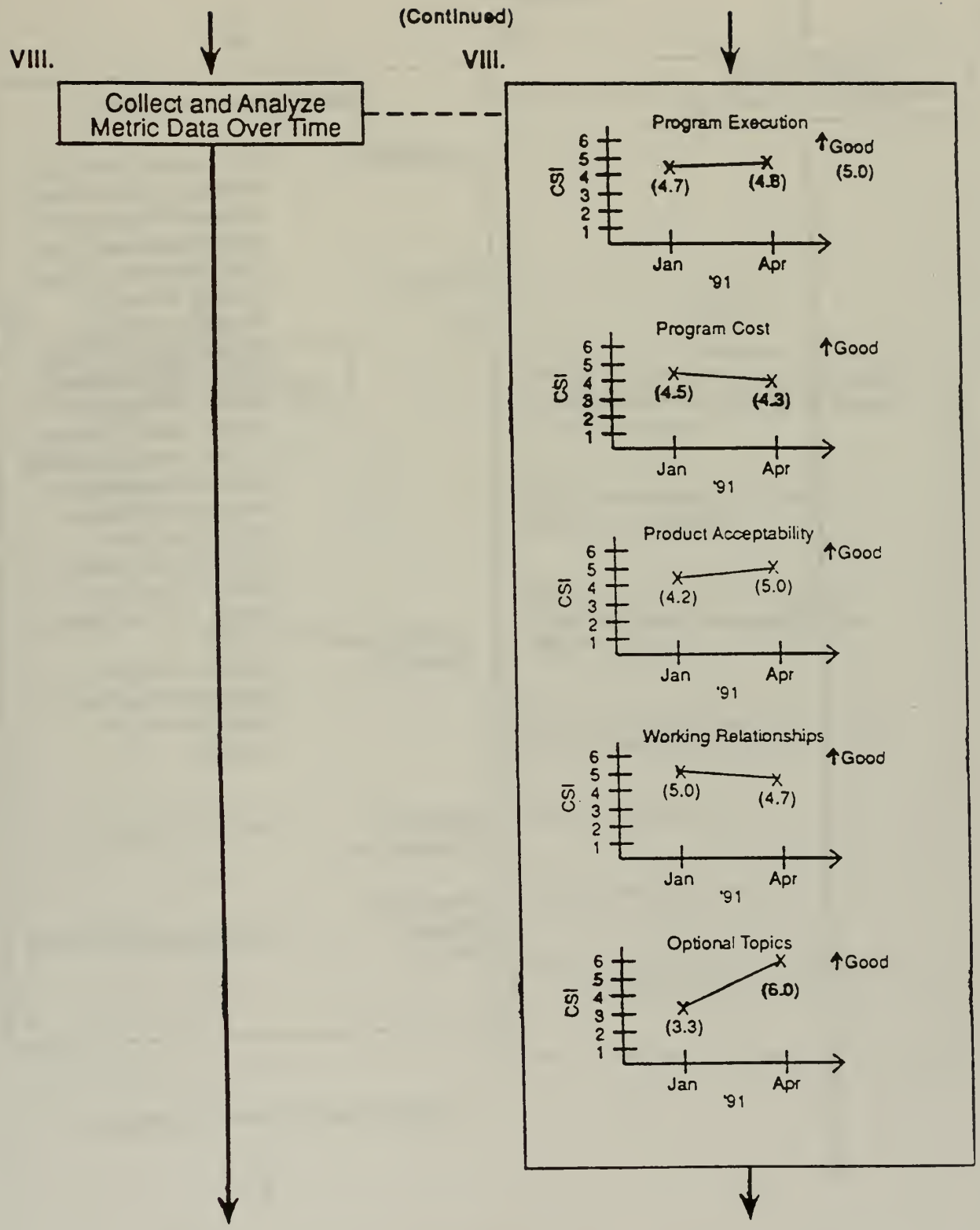

Figure 3-4 (cont). Metric Development Process. From METRICS, p. 3-8.

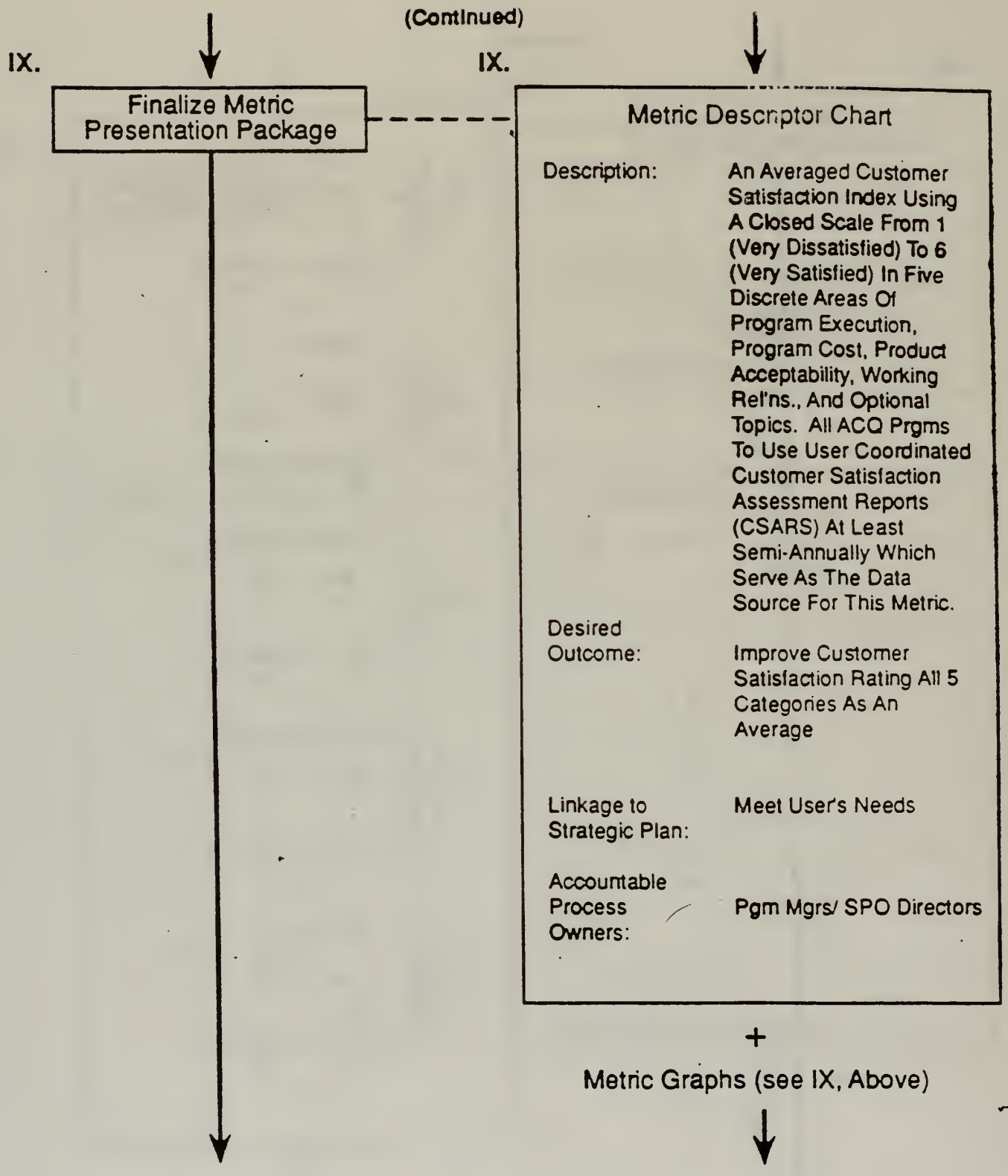
(Continued)
IX.
IX.
Finalize Metric Presentation Package
Metric Descriptor Chart
Description:
An Averaged Customer Satisfaction Index Using A Closed Scale From 1 (Very Dissatisfied) To 6 (Very Satisfied) In Five Discrete Areas Of Program Execution, Program Cost, Product Acceptability, Working Rel'ns., And Optional Topics. All ACQ Prgms To Use User Coordinated Customer Satisfaction Assessment Reports (CSARS) At Least Semi-Annually Which Serve As The Data Source For This Metric.
Desired Outcome:
Improve Customer Satisfaction Rating All 5 Categories As An Average
Linkage to Strategic Plan:
Meet User's Needs
Accountable Process Owners:
Pgm Mgrs/ SPO Directors
+
Metric Graphs (see IX, Above)

Figure 3-4 (cont). Metric Development Process. From METRICS, p. 3-9.

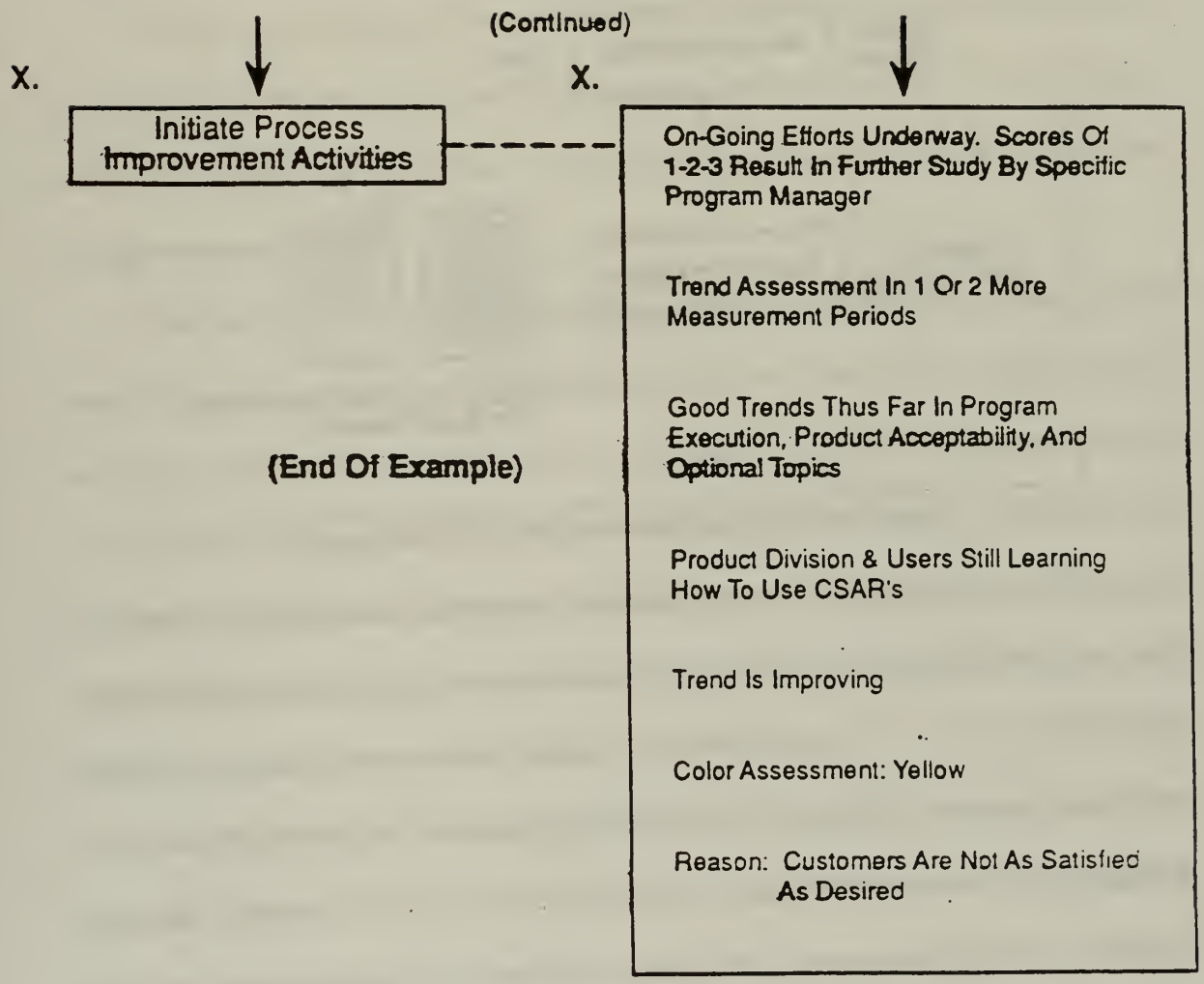
(Continued)
X.
X.
Initiate Process Improvement Activities
On-Going Efforts Underway. Scores Of 1-2-3 Result In Further Study By Specific Program Manager
Trend Assessment In 1 Or 2 More Measurement Periods
Good Trends Thus Far In Program Execution, Product Acceptability, And Optional Topics
(End Of Example)
Product Division & Users Still Learning How To Use CSAR's
Trend Is Improving
Color Assessment: Yellow
Reason: Customers Are Not As Satisfied As Desired

Figure 3-4 (cont). Metric Development Process. From METRICS, p. 3-10.

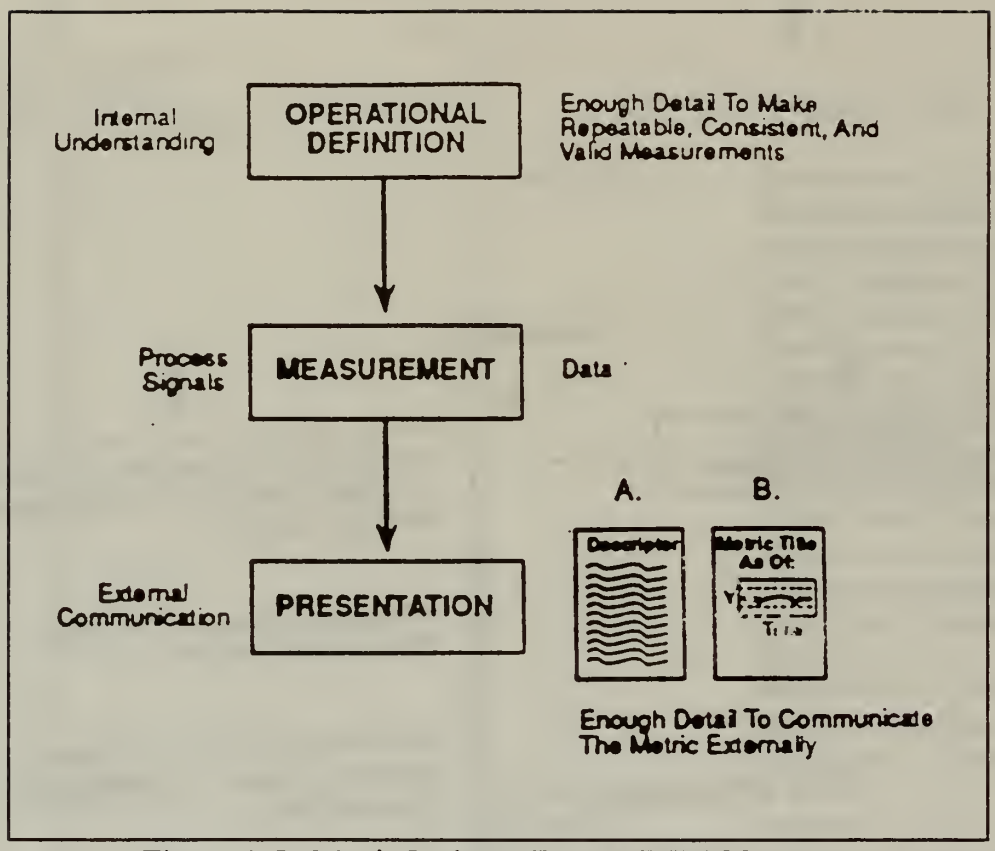

Figure 3-5. Metric Package. From METRICS, p. 3-1.

The second element of a metric package, the actual measurement and recording of data, "is the translation of data from the process into understandable and useful information" (METRICS, 1991, p. 3-1). And the third element of the metric package, the metric presentation, consists of two parts: "The first part is the metric descriptor and the second part is the graphic presentation of the data" (METRICS, 1991, p. 3-2). The "metric descriptor" is simply a written description that clearly expresses the information in the graphic presentation.

A fundamental attribute of any management tool is simplicity. The Metric Development Process lacks this attribute. Not only are the steps complicated and involved, the process itself is unwieldy -- its illustration is five pages long.

That criticism aside, it is still a highly useful tool for evaluating metrics and linking them to outcomes. The Air Force's process unquestionably captures the "how-to" of metric development and evaluation, which is relevant to this study's objectives.

C. APPROACH SELECTED

The ideal approach to performance measurement will meet the primary objectives of this study, and it will satisfy the modeling principles discussed in section A. As discussed earlier, it is feasible and desirable to apply Liao's modeling principles to the performance measurement methods.

The four approaches are all very good methods for establishing a measurement system. In many respects, they possess significant similarities. None is lacking in any material way, which makes it difficult to select a preferred method. The selection process will attempt to choose the most suitable approach based on a comparison of all four. Admittedly, this selection process allows broad leeway in determining which approach is most suitable. However, that drawback does not significantly limit its utility in evaluating the approaches, as professional judgment is often relied on in decision-making.

In terms of simplicity and understandability, no clear candidate emerges as the "best" approach. The Metric Development Process of the Air Force is disqualified in terms of simplicity because of its complexity in comparison to the other approaches. The remaining three approaches have about the same degree of simplicity. The TRADE Performance Measurement Process has a slight edge over the other methods because of its identified outputs at each step throughout the process. This advantage allows the user to be fully aware of what is supposed to happen at each step, and increases one's understanding of the overall

process.

The framework that best captures the essence of measurement is the TRADE Performance Measurement Process. In the first step, it calls for identification of the process flow. This is a key idea -- one must examine and understand the process to be measured before any meaningful measurement system can be put in place. According to TRADE, "before you try to control a process, you must understand it" (1997, p. 5). The other approaches are deficient in this area.

Another important criterion is linking metrics to an outcome. This essentially means tying a measure to some consequence or result. The Metric Development Process is clearly superior in this area. Its "metric package" is a step-by-step process for developing measures of effectiveness. A key element of the metric package is the operational definition of the metric which includes the desired outcome and "the link between the process being measured, your organization's strategic plan, and the command [organization's] goals" (METRICS, 1991, p. 3-2).

TRADE's Performance Measurement Process also emphasized this important connection. TRADE pointed this out by stating: "Perhaps the most fundamental step in establishing any measurement system is answering the question, 'What do I want to know.' The key issue then becomes, 'How do we generate useful information?' . . . It is crucial to be able to state precisely what it is you want to know about the activity you are going to measure. Without this knowledge, there is no basis for making measurements" (1997, p. 5).

The final criterion for selecting a preferred approach is identifying metrics that are accurate indicators of performance. Two approaches are strong in this respect: The

Performance Measurement Process (TRADE) and the Metric Development Process (Air Force). The Training Resources and Data Exchange group presented a functional way to categorize metrics, shown in Figure 2-2. This clearly showed the relationship between category of measurement (i.e., quality, efficiency, effectiveness, timeliness, etc.) and the type of metric needed to measure it. This notion helps to focus on the connection between the metric and what it measures.

Based on the evaluation factors, the preferred approach for this research is a combination of two approaches: The Performance Measurement Process and the Metric Development Process. This approach will be described in the next chapter.

D. CHAPTER SUMMARY

This chapter defined a performance measurement approach as a step-by-step procedure for converting measurement data into information to improve a process or activity. This type of framework is expected to be able to identify metrics that are accurate indicators of performance. If successful, the selected approach will link metrics with specific performance outcomes.

The four approaches looked at were the General Measurement Methodology (Sink and Tuttle), the TRADE Performance Measurement Process, the Measurement Linkage Model (Chang and De Young), and the Metric Development Process (U.S. Air Force). Strengths and weaknesses of each approach were identified and discussed.

Although not models, these four approaches can be evaluated on the same general criteria used to evaluate a model. These criteria are: It should be simple, accurate and understandable; it should be relevant to the study's objectives; and it should be an authentic

replication of the situation it represents. The preferred approach for this research was determined to be a composite of the TRADE Performance Measurement Process and the Metric Development Process. The best attributes of these two processes will be presented in Chapter IV.

IV. DATA ANALYSIS

A. INTRODUCTION

This chapter describes and depicts the preferred performance measurement approach for this research, and presents a template for metric development and evaluation. It then applies a recognized DAWIA metric, professional certification rates, to the metric template to validate performance measurement characteristics discovered by the research.

B. METRIC ASSESSMENT AND MEASUREMENT APPROACH

The preferred performance measurement approach to use as an analysis tool for DAWIA is a modified combination of the TRADE Performance Measurement Process and the Metric Development Process of the U.S. Air Force Systems Command; both processes were described and depicted in Chapter III. The best features of both approaches are joined to create a more useful method to evaluate DAWIA metrics and link them to some outcome.

In keeping with the military's fondness for acronyms, this tailored approach is named the Metric Assessment and Measurement Approach (MAMA). Its name is appropriate because it is both a metric assessment tool and a performance measurement approach. Key features of the Metric Assessment and Measurement Approach are:

- a team approach to performance measurement
- process flowcharting
- specified outputs at each step
- identification of personnel responsible for specific actions
- the Metric Template

- the Metric Package

The significance of each of these features was discussed in Chapter III, and will not be repeated here. The Metric Assessment and Measurement Approach is a complete "how-to" method for establishing a measurement system. It is important to note that before beginning the process, a measurement team must be established -- this is a preparatory step taken before using the Metric Assessment and Measurement Approach. The team is responsible for designing and implementing the measurement process. The importance of using a team cannot be overemphasized. Sink and Tuttle (1989) found that the team concept "ended up with high-quality measurement systems, greater acceptance of the results, and a better foundation for moving through to implementation of the measurement systems" (p. 232). For the reasons stated above, it is important to not overlook this key step in using any sophisticated performance measurement system.

The 11 steps of the Metric Assessment and Measurement Approach are described below, and are shown in Figure 4-1.

> *Step 1. Decide what to measure.* First determine which category (implementation, efficiency, effectiveness, etc.) to measure. Then figure out the critical activities (i.e., those activities that are strategically important to success) for that category. Output of step 1: Identification of the measurement category, and a list of critical activities for the process to be measured.
>
> *Step 2. Flowchart the process.* To be effective, the team needs a common understanding of the process to be measured. Output of step 2: A list of key processes and flow charts for each.
>
> *Step 3. Develop and evaluate metrics.* Generate metrics. Develop operational definitions and evaluate each. Use the metric template (Figure 4-2) to do this. Output of step 3: A list of metrics with a completed metric template for each.

Step 4. Establish performance standards. Metrics should be linked to a predefined criterion or standard. Output of step 4: A list of standards for each critical activity.

Step 5. Identify responsible parties. Decide who will gather, examine and report data, how it will be done, who will decide if corrective action is needed, and who will make those changes. These people may or may not be team members. Output of step 5: A list of people, their areas of responsibility, and data collection forms.

Step 6. Collect data. Using the metrics identified in step 3, gather data. Check for early trends to ensure the right data is being collected and is meaningful. Output of step 6: Completed data collection forms and/or checksheets.

Step 7. Analyze/report actual performance. Metric data are converted into a graphic presentation that form a basis for decision-making. Output of step 7: Data presentation in "metric package" form and/or a report.

Step 8. Compare actual performance to standards. Compare reported performance to predetermined performance standards and determine the variation (if any). Output of step 8: A decision based on performance variance.

Step 9. Determine if corrective actions are necessary. Step 9 is a decision step. You can either change the process to bring it up to standard; or modify the standard to make it more realistic. Output of step 9: An action plan to make changes or reevaluate standards.

Step 10. Make changes to bring process up to performance standard. This step only occurs if corrective action is necessary. Decide what action to take. Make changes. Output of step 10: A successfully implemented plan.

Step 11. Decide if new standards or new metrics are needed. Changes in the marketplace, customer needs, government regulations, etc. may require reevaluation of performance standards. Regardless, standards should be reviewed annually at a minimum. Output of step 11: New standards, new metrics, or no change.

The Metric Assessment and Measurement Approach follows a clear and logical process for converting measurement data into usable information for decision-making. It incorporates two measurement tools, the Metric Package and the Metric Template, which are described in the next sections.

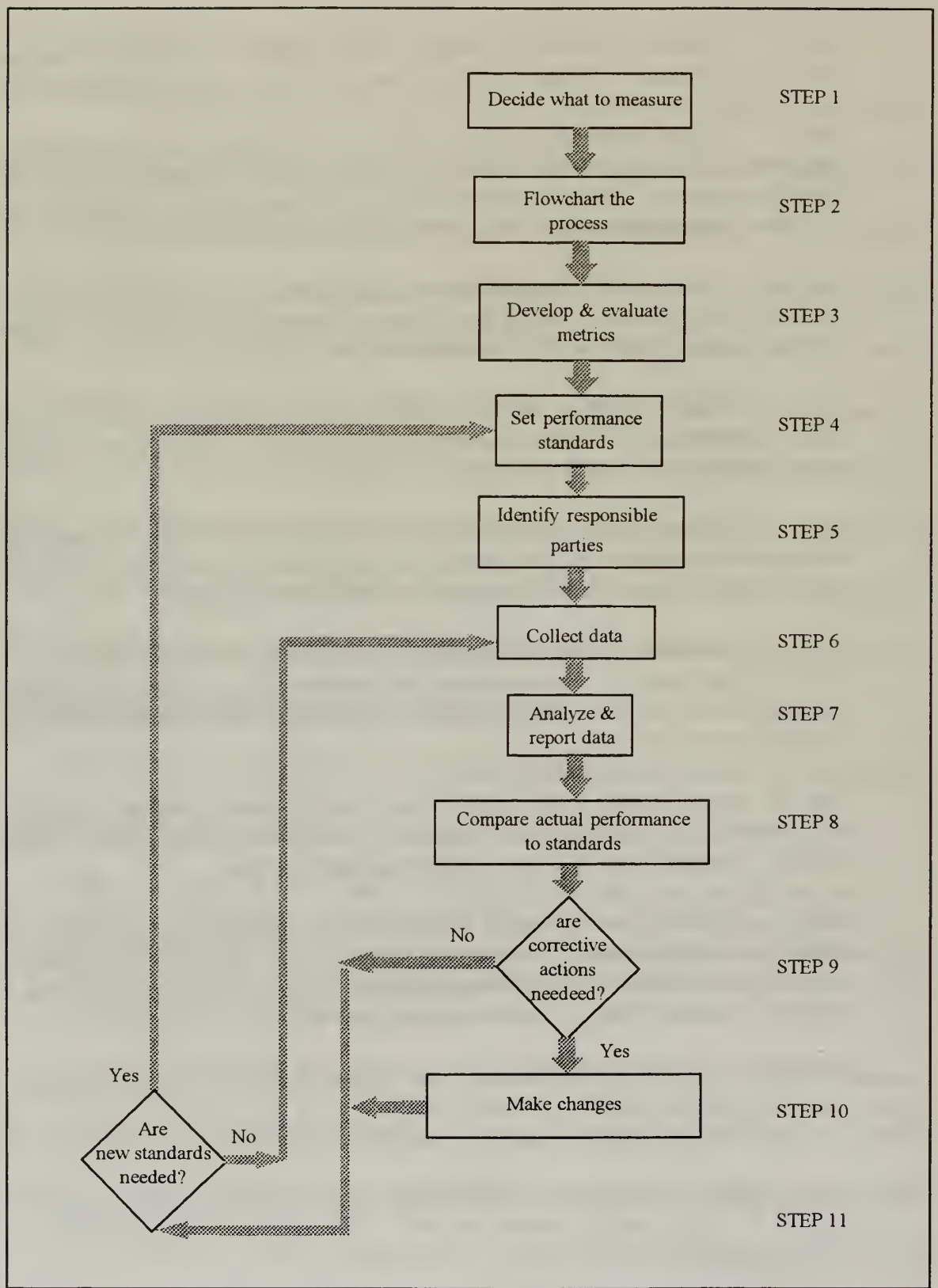

Figure 4-1. Metric Assessment and Measurement Approach.

1. Metric Template

In attempting to find an easy way to evaluate metrics, which occurs in step 3 of the Metric Assessment and Measurement Approach, a template was developed. The template serves several functions. It allows the user to:

- identify the metric category
- give the metric an operational definition
- evaluate the metric to ensure its usefulness
- identify the customer of the metric
- identify ratios or equations needed to gather metric data

The template is a simple tool to ensure quality metrics. It has three parts: identification, definition, and evaluation. The first part of the Metric Template (shown in Figure 4-2) is metric identification. This consists of selecting the metric category and naming the metric.

The second part is the operational definition. It consists of a description of the metric, a complete explanation of the process to be measured, and identification of the desired outcome in terms of what results the organization intends to achieve. The operational definition, as discussed in Chapter III, is needed to communicate an ambiguous concept, such as "reliability," for example. How is reliability measured? It has to be expressed in a way that can be counted, such as failure rates, or stated in more positive terms, number of successive periods without failures.

The third part of the metric template allows the user to evaluate the metric. The first column lists desirable metric attributes, and the second and third columns are used to simply

METRIC TEMPLATE

PART I. METRIC IDENTIFICATION

METRIC CATEGORY: (circle one) Effectiveness Efficiency Quality Timeliness Productivity Implementation Financial Results Safety

METRIC NAME:_____

PART II. OPERATIONAL DEFINITION

WHAT (The information in this section represents the link between the metric and an outcome.)

1. What does the organization want to know?:_____
2. Desired outcome:_____
3. Description of metric:_____
4. Precise explanation of the process to be measured:_____

WHO

1. Customer of the metric:_____

HOW

1. Ratios or equations to be used:_____
2. How metric data will be displayed:_____

PART III. METRIC EVALUATION

METRIC ATTRIBUTE	YES	NO	DESCRIBE HOW ATTRIBUTE APPLIES
*meaningful to customer and organization			
*relates to organizational goals			
*enables decision making			
*metric data is measurable			
*measured process can be controlled			
metric data is economical to collect			
metric data is timely			
compatible with existing measurement system			
simple, understandable, and repeatable			
enables trend analysis			
clearly defined			

***IF THE ANSWER IS "NO" TO ANY OF THESE KEY ATTRIBUTES, DISCARD THE METRIC!**

Figure 4-2. Metric Template.

check off whether the metric satisfies the attribute. The fourth column provides space for a short description of how the metric satisfies the attribute. If the user is unable to articulate how the attribute relates to the metric in two or three instances, it may not be a useful metric, and probably should be discarded.

By using the metric template, the metrics are formalized as useful indicators of performance. In addition to many of the metric attributes identified in Chapter II, two measurement concerns have been satisfied in the template: validity and reliability. Validity is established when the selected measure is related to the event or object that can be controlled or manipulated. This criterion is met if the organization can successfully identify what it wants to know (question 3, Part II. of the metric template), and it can control some process relating to what it wants to know. Reliability was defined, in the context of measurement, as consistency and accuracy. If the metric data can be gathered repeatedly, under the same conditions, then it is reliable.

2. Metric Package

Step 7 of the Metric Assessment and Measurement Approach calls for a "metric package," which is borrowed from the Metric Development Process of U.S. Air Force Systems Command. Although this was shown and discussed in Chapter III, more detail will be provided here.

The two elements of a metric package are a metric descriptor and the graphical presentation of metric data. Figure 4-3 shows the format to be used for a metric descriptor. A clear and concise written description of the metric will immediately convey what was measured and how it was measured. The resulting measurement data is what will be

Metric Title

Description:
Briefly define your metric along with the population to be measured and the source of your data. These and other items of information on your metric should be contained in your metric operational definition.

Desired Outcome:
Define the outcome in terms of improved processes. Do not use numerical goals.

Linkage to Strategic Plan:
Identify one or more of your organization's objectives or goals that are addressed by tracking this metric. This linkage to your organization's business plan is essential; metrics for metrics' sake are unwarranted.

Process Owner:
Identify the principal individual or organization who can initiate and sustain process improvement.

Figure 4-3. Metric Descriptor. From METRICS, p. 5-1.

displayed in the graphical presentation.

Figure 4-4 depicts graphical presentation of metric data for the DAWIA professional certification metric. Four things are shown in this chart:

- number of personnel whose qualifications exceed their required certification level (shaded area at the bottom of the bar chart)
- number of personnel whose qualifications meet their required certification level (darkly shaded area)
- number of personnel whose qualifications do not meet their required certification level (white area)
- "other" (lightly shaded area at top)

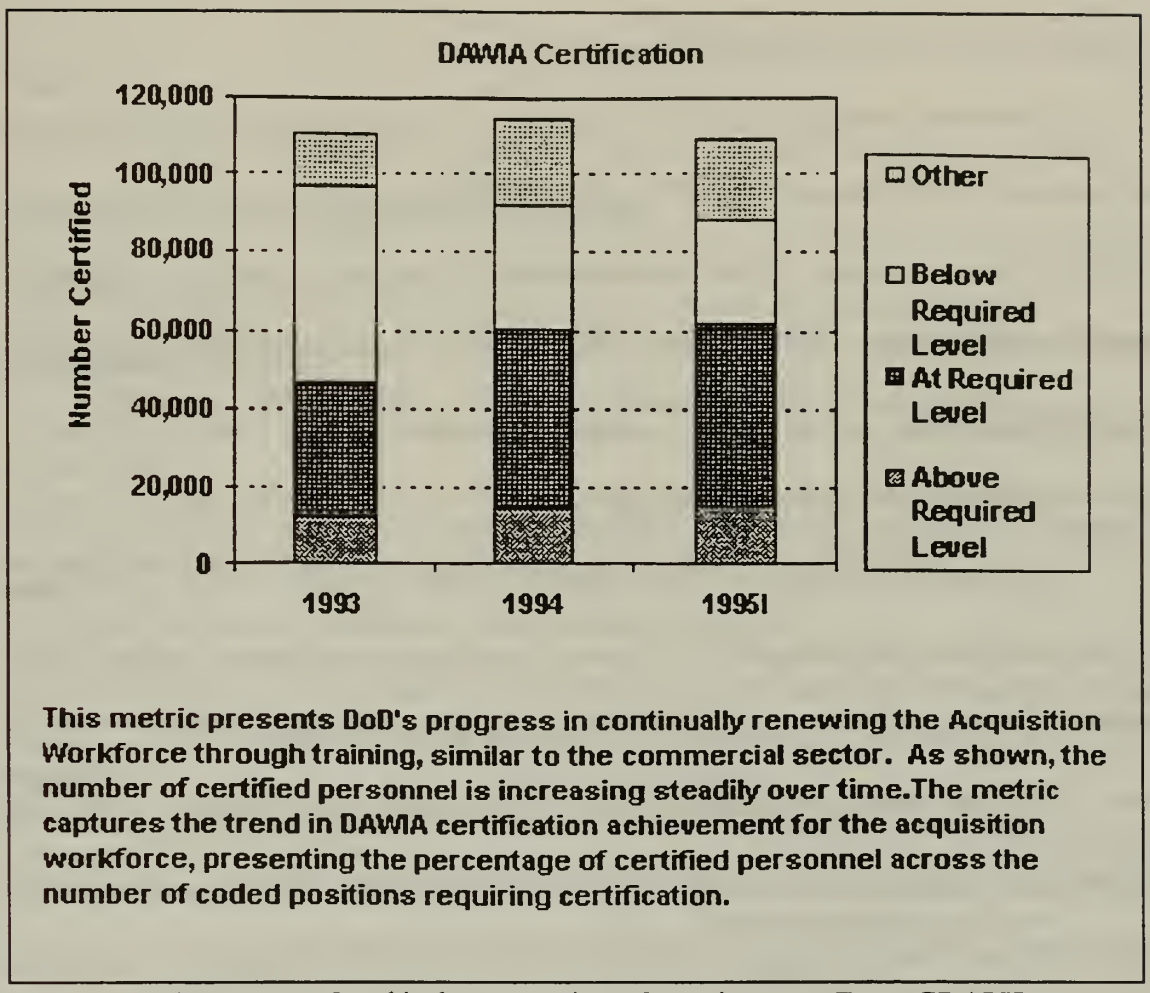
DAWIA Certification
120,000
100,000
80,000
60,000
40,000
20,000
0
Number Certified
1993
1994
1995
Other
Below Required Level
At Required Level
Above Required Level
This metric presents DoD's progress in continually renewing the Acquisition Workforce through training, similar to the commercial sector. As shown, the number of certified personnel is increasing steadily over time.The metric captures the trend in DAWIA certification achievement for the acquisition workforce, presenting the percentage of certified personnel across the number of coded positions requiring certification.

Figure 4-4. Graphical Presentation of Metric Data. From GRAPH.

Below the chart is an inaccurate description of what is depicted. The description says it represents "the percentage of certified personnel across the number of coded positions requiring certification" (GRAPH). What the chart actually shows is the number of certified personnel, not the percentage of certified personnel. The white area shown in the bar charts is perhaps the most useful piece of information. It shows a trend over the three-year period toward fewer workers who have not met their required certification level. Conversely, the bottom two areas together in each bar show growth over time in the number of personnel who have met or exceeded their required certification level. This would be the preferable

measure to track workforce certification rates.

A weakness in this particular metric presentation (Figure 4-4) is the lightly shaded area at the top of each bar labeled "other." There is no description of what it represents. This is pointed out as a weakness in this metric presentation. All data, scales, and x- and y-axes should be clearly labeled. The data will not be useful if misrepresented, or if misunderstood by a decision-maker.

C. VALIDATION OF APPROACH

In this section, the only DoD DAWIA metric currently in use, professional certification rates, will be applied to the metric template for analysis and validation. This metric will be applied to the template as an "implementation" category metric.

Figure 4-5 shows a completed template for DAWIA certification rates. The first part is relatively straightforward; it needs no additional explanation. Part 2 provides the link between the metric and an outcome. According to TRADE (1997), "In the end, what is done and measured somehow must connect with the desired outcome" (p. 16). To achieve this connection, first state what you want to know about some process; then identify outcomes you want to verify with a metric. Check to see if the metric you have chosen will provide that information. If not, find another metric.

In the certification rate example, the organization wants to know if certification of personnel has taken place, as required by DAWIA. The desired outcome is 100 percent certification of the workforce by a specified date. The metric is defined as the ratio between the number of defense acquisition personnel certified and total number of defense acquisition personnel. Is this metric linked to an outcome? It is clearly linked to its desired outcome.

METRIC TEMPLATE

PART I. METRIC IDENTIFICATION

METRIC CATEGORY: (circle one) Effectiveness Efficiency Quality Timeliness Productivity **Implementation** Financial Results Safety

METRIC NAME: **DAWIA Professional Certification Rate**

PART II. OPERATIONAL DEFINITION

WHAT (The information in this section represents the link between the metric and an outcome.)

1. What does the organization want to know?: **is the organization in compliance with DAWIA requirement to get the workforce certified?**
2. Desired outcome: **100% certification of defense acquisition personnel by Sept 30, 1998**
3. Description of metric: **percentage of defense acquisition workforce that have met professional certification requirements**
4. Precise explanation of process to be measured: **count the number of workforce who have met certification requirements; count total number of defense acquisition personnel; divide first figure by second figure, and the result is percentage (expressed in decimal form) of workforce that are certified.**

WHO

1. Customer of the metric: **Under Secretary of Defense for Acquisition**

HOW

1. Ratios or equations to be used: **number certified defense acq. personnel/number defense acq. personnel**
2. How metric data will be displayed: **pie chart or bar chart**

PART III. METRIC EVALUATION

METRIC ATTRIBUTE	YES	NO	DESCRIBE HOW ATTRIBUTE APPLIES
*meaningful to customer and organization	X		**USD (Acquisition) must ensure professional qualification of the workforce**
*relates to organizational goals	X		**required by law**
*enables decision making	X		**shows how training resources should be used to ensure all personnel obtain training**
*metric data is measurable	X		
*measured process can be controlled	X		
metric data is economical to collect	X		**can be reported by command or activity**
metric data is timely	X		**can be collected annually, semi-annually, quarterly, etc.**
compatible with existing measurement system	X		**data is available in personnel training records**
simple, understandable, and repeatable	X		
enables trend analysis	X		**period comparisons are feasible**
clearly defined	X		

***IF THE ANSWER IS "NO" TO ANY OF THESE KEY ATTRIBUTES, DISCARD THE METRIC!**

Figure 4-5. Metric Template using a sample "Implementation" Metric.

For certification rate, each attribute in Part 3 of the template is adequately satisfied. A brief description has been provided in the last column to explain the attribute's relevance to the metric. The first five attributes are vital to ensuring a valid measurement process. The asterisk indicates that if any one of these five attributes are not satisfied, the metric must be rejected. A metric was defined first and foremost as a meaningful measure. It should be meaningful to both the customer and the organization. Equally important, "it must present data that allow us to take action" (METRICS, 1991, p. 1-1). In other words, the metric provides information that enables decision-making.

Another critical metric characteristic is whether the process for which the data is collected can be controlled. If it cannot be controlled or manipulated in some way, measuring it serves no purpose, since nothing can be done to change or improve the process. For the example shown in Figure 4-5, the process that is being measured and *can be controlled* is professional certification rates. Obviously, something can be done to improve workforce certification, such as increase training opportunities for the workforce.

Lastly, the metric must relate to organizational goals. If it does not, the metric data will have no impact on the organization's objectives. The DAWIA certification rate metric met all five of these critical attributes, and all of the other attributes listed in Part 3.

D. CHAPTER SUMMARY

This chapter presented a performance measurement approach suitable for evaluating the effectiveness of DAWIA. The approach is a blend of the TRADE Performance Measurement Process and Air Force System Command's Metric Development Process. The new approach is called the Metric Assessment and Measurement Approach. Its primary

advantages are:

- a team approach
- flowcharting the process to be measured
- specified outputs at each step
- identification of personnel responsible for specific actions
- it establishes linkage between the metric and an outcome

This chapter also introduced a metric template for creating, operationalizing, and evaluating metrics. The template was tested and validated with the only existing DoD DAWIA metric, professional certification rate. This metric was analyzed against the Metric Assessment and Measurement Approach and the research literature. The metric successfully met attributes described in the performance measurement literature.

The next chapter presents conclusions and recommendations drawn from this research.

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

This research discussed the use of program measures to assess outcomes of public policy. Each year new laws are enacted with the rightful expectation that they will achieve desired results. Unfortunately, this is not always the case. The 1990 Defense Acquisition Workforce Improvement Act is a public law designed to improve acquisition practices in the Department of Defense through management of the acquisition workforce. This Act was looked at in terms of how to develop and evaluate metrics that will be good indicators of acquisition workforce performance. The following conclusions are made:

- *Training and education of the defense acquisition workforce are instrumental to improving performance.* This finding was supported by numerous defense acquisition studies over a period of 35 years. Yet, professional training and career development of the workforce was never formally implemented until enactment of DAWIA.

- *Performance measurement is needed to understand the impacts of public policies, laws and federal programs.* Actual consequences of a policy or program must be compared to its original objectives using some type of measurement/assessment tool to find out if it has achieved its goals. The 1993 Government Performance and Results Act makes performance measurement mandatory for all federal agencies.

- *For measurement data to be useful and effective, measures must be valid, reliable, and unbiased.* Brinkerhoff and Dressler (1990) explained that validity establishes the link between what is measured and what the person doing the measurement wants to know. Reliability refers to the measure's consistency and accuracy, while bias is a defect in the data collection or compilation process.

- *Nonjudgmental measures are preferable to judgmental measures.* Nonjudgmental measures are based on tangible data and are objective. Judgmental measures deal in abstractions and much more susceptible to bias.

- *Dysfunctional measurement paradigms can compromise a measurement system.* Two paradigms are particularly damaging: the use of a single measure to assess performance, and workers' fear that measurement data will be used against them.

- *The Metric Assessment and Measurement Approach is the preferred performance measurement approach for DAWIA.* It is a combination of the TRADE Performance Measurement Process and the U.S. Air Force System Command's Metric Development Process. The Metric Assessment and Measurement Approach uses a team strategy. This approach requires process flowcharts and specifies an output at each step. It also prompts the user to identify specific personnel and their measurement responsibilities and establishes a link between metrics and desired outcomes.

- *The metric template is an effective tool for creating, operationalizing, and evaluating metrics.* The template enables the user to identify metric category, to operationalize and evaluate the metric, and to identify the metric's customer. It also facilitates the development of ratios or equations to be used to gather and analyze metric data. The template was successfully used to validate the only DAWIA metric currently in use, professional certification rates.

B. RESEARCH QUESTIONS ANSWERED

The primary research question was: Which metrics are good indicators of acquisition workforce performance? In answering this question, one must recognize the different categories or classifications by which performance can be measured. Those categories include effectiveness, efficiency, quality, timeliness, productivity, financial results and safety. The metric will depend on what information is being sought about a process or activity. Completing the metric template (Figure 4-2) will enable the user to determine if the metric is a good indicator of workforce performance.

Two interrelated research questions were: How can suitable metrics for workforce performance be identified, and can a metric template be developed that captures performance indicators? The first question is answered with the metric template that was developed in response to the second question. Many metric attributes were identified by METRICS

(1991), TRADE (1997), and others. Several of the attributes they identified are inherent to good quality measurement systems. It makes sense that metrics must enable the collection of data that can provide information for decision-making. Ultimately, the purpose for taking measurements is to use the measurement data to make decisions. Therefore, many of the metric attributes in the template are tied to providing information for decision-making.

Two final research questions were: How can metrics be evaluated as good indicators using a performance measurement approach, and How can acquisition workforce metrics be linked to performance improvement outcomes? The metric template was used to resolve both of these questions. Metric evaluation occurs during Step 3 of the Metric Assessment and Measurement Approach by leading the user to the metric template. Part 3 of the template is for metric evaluation. The metric is judged to be effective if it meets each attribute. Part 2 of the template provides the link between a metric and an outcome. Part 2 "walks" the user through the process of identifying a desired outcome, how the measurements will be conducted, and what ratios will be used. Putting everything down in writing on the metric template enables the user to think through the measurement process, and get a clear picture of what information is needed, how it will be obtained, and how it will be used.

C. RECOMMENDATIONS DRAWN FROM RESEARCH

The first recommendation is that the Metric Assessment and Measurement Approach and metric template be put to use by any public sector organization that needs to know how its outcomes compare to its original objectives.

Another recommendation concerns the sharing of performance measurement information, techniques and results. This type of information is valuable to virtually any

organization -- public sector, private sector and nonprofit. Recommend defense activities and other federal government agencies post useful measurement information and insights on the Internet World Wide Web. With continued cuts in government programs, such as defense acquisition programs, it is necessary to share success with other agencies. Two suggested Internet performance measurement websites are: http://www.itpolicy.gsa.gov/mkm/pathways/pathways.htm, and http://www.dtic.mil/c3i/c3ia/itprmlinks.html.

D. RECOMMENDATIONS FOR FURTHER RESEARCH

Other areas of metric development and evaluation for DAWIA remain to be explored. One of DAWIA's objectives is to train and educate acquisition personnel, which should result in a more effective workforce. Defense Acquisition University (DAU) provides this job-related education and training. But do DAU training programs actually result in a more effective workforce? We know that the output of training is some number of people who receive a certificate of completion. What is important is not the *output* of training, but the *outcome* of training. Is the workforce more effective after being properly trained? How is training effectiveness measured? These questions and others could be examined in depth.

During this research, only a single metric (professional certification rates) was in use to measure DAWIA. The Metric Assessment and Measurement Approach or some similar approach should be applied to DAWIA to determine if DAWIA achieved its desired outcomes. More DAWIA metrics must be developed to answer this question. Pope (1997) suggested numerous metrics for DAWIA. His metrics should be evaluated using the metric template and put into use if they are shown to be valid, reliable, unbiased, and meet the other key metric attributes.

The Government Performance and Results Act requires all federal agencies[5] to develop a strategic plan, set goals, establish performance indicators, and measure and report performance results annually to Congress. The requirements of the GPRA should be examined to look for ways it can be applied to DAWIA. Also, the results of GPRA's pilot projects should be reviewed for "lessons learned" and to gain an understanding other agencies' measurement experiences. A recent GAO report (see GAO (ii) reference) has published early results of selected regulatory agencies in complying with GPRA.

Performance measurement provides a reliable foundation for decision-making. It is important for the Department of Defense and other federal agencies to implement ongoing performance measurement programs to monitor, control and improve their processes, and ultimately their outcomes. The Defense Acquisition Workforce Improvement Act was intended to result in a better trained and more effective acquisition workforce. Better workers would, in theory, improve military procurement and save taxpayer dollars. Policymakers, taxpayers, and other stakeholders are entitled to some assurance that DAWIA has achieved, or is making progress toward achieving its intended outcomes.

---

[5] Five government agencies are exempt from the GRPA: Central Intelligence Agency, General Accounting Office, Panama Canal Commission, Postal Service, and Postal Rate Commission.

APPENDIX. GLOSSARY OF TERMS

Acquisition: The conceptualization, initiation, design, development, test, contracting, production, deployment, and logistic support, modification, and disposal of weapons and other systems, supplies, or services (including construction) to satisfy DoD needs, intended for use in or in support of military missions. (Glossary, 1995, p. B-1)

Acquisition Corps: Highly qualified acquisition professionals in the grades of Navy Lieutenant Commander, Army or Marine Corps Major, and civilian personnel in grades GS-13/GM-13 and above.

Acquisition Workforce: The personnel component of the acquisition system. The acquisition workforce includes permanent civilian employees and military members who occupy acquisition positions, who are members of an Acquisition Corps, or who are in acquisition development programs. (DoD Directive Number 5000.52, 1991, enclosure 2)

Critical Acquisition Position: Those senior positions carrying significant responsibility, primarily involving supervisory or management duties, in the DoD acquisition system. (DoD Directive Number 5000.52, 1991, enclosure 2)

Defense Acquisition System: A single uniform system whereby all equipment, facilities, and services are planned, designed, developed, acquired, maintained, and disposed of within the DoD. (Glossary, 1995, p. B-30)

Defense Acquisition University (DAU): A consortium of DoD educational institutions that provide acquisition courses for the defense acquisition workforce.

Defense Acquisition Workforce Improvement Act (DAWIA): Legislation passed by Congress in 1990 which mandates training, education, and professional qualifications for persons serving in acquisition positions in the Department of Defense.

Government Performance and Results Act (GPRA): Results-oriented legislation enacted by Congress in 1993 which requires federal agencies to develop a five-year strategic plan, set performance goals, identify associated performance measures, and submit annual progress reports to Congress.

Program Management: The process whereby a single leader exercise centralized authority and responsibiliity for planning, organizing, staffing, controlling, and leading the coombined efforts of participating/assigned civilian and military personnel and organizations, for the management of a specific defense acquisition program or programs, through development, production, deployment, operations, support, and disposal. (Schmoll, 1996, p. 61)

Program Manager (PM): Official responsible for managing a specific acquisition program . . . . (Glossary, 1995, p. B-84)

Metric: A meaningful measure that provides useful data for decision-making.

Metric Template: A standardized format for defining, describing, and evaluating a metric, and linking it to an outcome.

Outcome: Consequences of an organization's activities.

Procurement: Act of buying goods and services for the Government (Glossary, 1995, p. B-78)

LIST OF REFERENCES

ACT: *Discussion of the acquisition workforce improvement act*. Internet World Wide Web address, http://www.dtic.dla.mil/acqed2/legislation/dawiadis.html (September 5, 1997).

Brinkerhoff, R. & Dressler, D. (1990). *Productivity measurement, a guide for managers and evaluators*. Newbury Park, CA: Sage Publications.

Cave, M., Kogan, M. & Smith, R. (eds.) (1990). *Output and performance measurement in government, the state of the art*. London: Jessica Kingsley Publishers.

Chang, R. & De Young, P. (1995). *Measuring organizational improvement impact*. Irvine, CA: Richard Chang Associates.

COMMISSION (i): Commission on Organization of the Executive Branch of the Government (May 1955). *Progress report as of December 31, 1954: A report to the congress*. Washington, DC: GPO.

COMMISSION (ii): Commission on Organization of the Executive Branch of the Government (June 1955). *Business organization of the department of defense: A report to the congress*. Washington, DC: GPO.

Deming, W. (1982). *Out of the crisis*. Cambridge, MA: MIT.

Dye, T. (1995). *Understanding public policy*. Englewood Cliffs, NJ: Prentice-Hall.

Franklin, J. & Thrasher, J. (1976). *An introduction to program evaluation*. New York: John Wiley & Sons.

Fremgen, J. (January 1997). *An introduction to internal control and auditing*. Unpublished manuscript. Naval Postgraduate School, Monterey, CA.

GAO (i): United States General Accounting Office. *Acquisition management: Implementation of the defense acquisition workforce improvement act* (April 1993). Gaithersburg, MD: GAO.

GAO (ii): United States General Accounting Office. *Managing for results: Regulatory agencies identified significant barriers to focusing on results* (June 1997). Washington, DC: GAO.

Glossary: Defense acquisition acronyms & terms (March 1995). Fort Belvoir, VA: Defense Systems Management College.

GPRA: *Government performance and results act of 1993*. Internet World Wide Web address, http://www.dtic.dla.mil/labman/gpra/law103-62.html (December 5, 1997).

GRAPH: *Enterprise metrics*. Internet World Wide Web address, http://www.acq.osd.mil/ar/arms/enterp05.htm (November 16, 1997).

HEARINGS (i): Hearings Before the Investigations Subcommittee of the Committee on Armed Services. House of Representatives, 101st Congress, First Session. *Department of Defense Implementation of the Packard Commission Report of 1986* (1990). Washington, DC: GPO.

HEARINGS (ii): Hearings Before the Investigations Subcommittee of the Committee on Armed Services. House of Representatives, 101st Congress, Second Session. *Acquisition workforce* (1990). Washington, DC: GPO.

Landy, F. & Farr, J. (1983). *The measurement of work performance: methods, theory and applications*. San Diego: Academic Press.

Liao, S. (n.d.). Unpublished manuscript. Naval Postgraduate School, Monterey, CA.

Manning, M. & Binzagr, G. (July 1996). Methods, values, and assumptions underlying large group interventions intended to change whole systems. *International Journal of Organizational Analysis, July 1996,* 1-13.

METRICS: U.S. Air Force Systems Command. *The metrics handbook* (August 1991). Andrews Air Force Base: HQ Air Force Systems Command.

Newbold, P. (1995). *Statistics for business & economics*. Englewood Cliffs, NJ: Prentice-Hall.

OECD: Organisation of Economic Co-operation and Development. *Performance measurement*. Internet World Wide Web address, http://www.oecd.org/puma/mgmtres/pac/key.htm (September 4, 1997).

PACKARD: United States President's Blue Ribbon Commission on Defense Management. *A formula for action: A report to the president on defense acquisition* (April 1986). Washington, DC: GPO.

Pope, J. (1997). *Measuring the effect of the defense acquisition workforce improvement act*. Unpublished research report, Naval Postgraduate School, Monterey, CA.

Render B. & Stair, R. (1997). *Quantitative analysis for management*. Upper Saddle River, NJ: Prentice Hall.

Rhoads, D. (1995). Is DAWIA worth it? An approach to analyzing the impacts. *Acquisition Review Quarterly*, Spring 1995.

Rummler, G. & Brache, A. (1990). *Improving performance: How to manage the white space on the organization chart*. San Francisco: Jossey-Bass.

Schmoll, J. (June 1996). *Introduction to defense acquisition management*. Fort Belvoir, VA: Defense Systems Management College.

Scioli, Frank P. and Thomas J. Cook (1975). *Methodologies for analyzing public policies*. Lexington, MA: D.C. Heath & Company.

Sink, D. and Tuttle, T. C. (1989). *Planning and measuring in your organization of the future*. Norcross, GA: Institute of Industrial Engineers.

Sink, D. (1985). *Productivity management: planning, measurement and evaluation, control and improvement*. New York: John Wiley & Sons.

TRADE: Training Resources and Data Exchange. *How to measure performance: A handbook of techniques and tools*. Internet World Wide Web address, http://www.llnl.gov/PBM/handbook (September 11, 1997).

United States. Department of Defense. *Defense acquisition education, training, and career development program*. Directive Number 5000.52, enclosure 2. (October 15, 1991).

BIBLIOGRAPHY

Brancato, C. (1995). *New corporate performance measures, a research report*. New York: The Conference Board.

Cook, T. & Scioli, F. A research strategy for analyzing the impacts of public policy. *Administrative Science Quarterly, September 17, 1972*, p. 328-338.

DAWIA: United States Congress, Public Law 101-50, Title XII (Nov. 1990). *Defense acquisition workforce improvement act*. Washington, DC: GPO.

OECD. Organisation of Economic Co-operation and Development (1994). *Performance measurement in government: issues & illustrations*. Paris: OECD.

Scioli, F. & Cook, T. (1975). *Methodologies for analyzing public policies*. Chicago: D.C. Heath & Company.

INITIAL DISTRIBUTION LIST

No. Copies

1. Defense Technical Information Center 2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5000

3. Professor Walter Owen (SM/On) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

4. Dr. Shu Liao (SM/Lc) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

5. LCDR Kathy E. Gordon, USN. 1
   7499 W. 600 N.
   Earl Park, IN 47942

6. Mr. & Mrs. Gale C. Gordon 1
   7499 W. 600 N.
   Earl Park, IN 47942

7. Maj. Chris Russick, USAF, Ret. 1
   640 Hillcrest Drive
   Fairborn, OH 45324

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

DUDLEY KNOX LIBRARY
3 2768 00342103 3